Kurmyshev

# Fundamentos de Métodos Matemáticos para Física e Ingeniería

LIMUSA
NORIEGA EDITORES

# Fundamentos de Métodos Matemáticos para Física e Ingeniería




# Fundamentos de Métodos Matemáticos para Física e Ingeniería

Evguenii V. Kurmyshev
Raúl E. Sánchez-Yáñez

Centro de Investigaciones en Óptica, A.C.





LIMUSA
NORIEGA EDITORES
MÉXICO • España • Venezuela • Colombia







GRUPO NORIEGA EDITORES
BALDERAS 95, MÉXICO, D.F.
C.P. 06040
(5) 8503-80-50
01(800) 7-06-91-00
(5) 512-29-03
limusa@noriega.com.mx
www.noriega.com.mx

CANIEM NÚM. 121

PRIMERA EDICIÓN
HECHO EN MÉXICO
ISBN 968-18-6366-6




# Contenido

# Prefacio

La complejidad de los problemas y sistemas a tratar en las ciencias exactas y naturales requiere de manera indispensable la aplicación de conceptos y métodos de las Matemáticas cada vez más complicados y poderosos. La aplicación de las Matemáticas, una herramienta de investigación científica bella y poderosa, a un problema de la Física o la Ingeniería consiste principalmente en tres fases de igual importancia:

**1. Modelo.** La elaboración de un modelo matemático consiste en presentar la relación entre los parámetros sustanciales (que se consideran como variables dependientes e independientes) de un proceso, sistema o fenómeno físico o de ingeniería en la forma de un sistema de ecuaciones matemáticas.

**2. Resolución.** El tratamiento del modelo por medio de métodos matemáticos nos lleva a una solución del problema dado en forma de una relación funcional entre variables dependientes e independientes del modelo.

**3. Interpretación.** La interpretación de la solución matemática en términos físicos que permite describir el comportamiento del sistema o fenómeno físico o de ingeniería bajo consideración.

Para lograr el éxito en cada una de las tres fases se requiere de intuición, de una forma de pensar matemática y del manejo de los métodos matemáticos. Sin embargo, en nuestros tiempos de "computación total", existe la creencia bastante común de que los problemas matemáticos en su mayoría se resuelven por computadora, es decir, a menudo se desprecia el conocimiento de los conceptos y métodos generales de las Matemáticas que se confunden con la habilidad de manejar, casi a ciegas, los programas computacionales ya hechos para las aplicaciones particulares. Es, por tanto, el objetivo y propósito principal de este libro, que el estudiante se familiarice profundamente con los conceptos matemáticos.

El motivo particular para la elaboración de los "Fundamentos de Métodos Matemáticos" surgió principalmente debido a la siguiente razón. La experiencia de investigador y docente de uno de los autores (E. Kurmyshev) a nivel de posgrado en física en varias materias (Métodos Matemáticos, Mecánica Cuántica, Mecánica Estadística, etc.), junto con las pláticas entre los colegas de trabajo y estudiantes, demostraron varias deficiencias e incongruencias que se mencionan a continuación. Ante todo, hay que reconocer que una parte de los estudiantes de posgrado son ingenieros por formación que en realidad tienen un nivel de estudios en Matemáticas un tanto bajo y conocimientos fragmentarios en la materia. Sin embargo, los programas de estudios en doctorado están elaborados

con base en la suposición de un nivel bastante alto que podría permitir los estudios a fondo de materias como mecánica cuántica, electromagnetismo, óptica, estado sólido y mecánica estadística que requieren métodos matemáticos avanzados. Además, varias veces me encontré con una situación un tanto extraña pero real, que dentro de un programa de estudios de doctorado, los métodos matemáticos se imparten en uno o dos semestres y, a veces, al mismo tiempo o después de las materias que los requieren. A pesar de que existen varios libros de texto excelentes en matemáticas y con una amplia cobertura de temas, la mayoría de éstos están escritos en inglés o algún otro idioma extranjero que, por supuesto, dificulta el aprendizaje a los estudiantes de habla española y aún más dentro de un tiempo muy limitado. Por tanto, surgió la necesidad de corregir las incongruencias existentes.

Como requisitos para escribir los "Fundamentos de Métodos Matemáticos" nos ajustamos a los siguientes lineamientos: 1) la selección de temas está apegada al programa de estudios de Doctorado en Física (Óptica) del Centro de Investigaciones en Óptica, A.C., y cubre las necesidades fundamentales en métodos matemáticos de varias materias como óptica clásica, mecánica clásica, electromagnetismo, las bases de teoría de ondas y oscilaciones, mecánica cuántica y mecánica estadística; 2) los "Fundamentos de Métodos Matemáticos" deben ser compactos, bastante rigurosos y con alta cultura matemática; 3) están dirigidos a los estudiantes del programa de posgrado en Física, pero pueden usarse como material de apoyo en ramas afines de ingeniería. El libro se pensó como autoconsistente y autosuficiente a su nivel.

El texto pretende abarcar los temas en matemáticas más indispensables hasta para estudios de Doctorado en Física (Óptica) y áreas afines. Dado que a la materia de matemáticas se le dedica un solo semestre en muchos de los programas de estudios de Doctorado en Física, sólo cambiando la selección de temas y su nivel es prudente ofrecer a los estudiantes un libro de texto que sea consistente y conciso a la vez, dejando una parte del desarrollo matemático para los estudiantes y no sobrecargando el texto con ejemplos. El material de un curso matemático demasiado *"machacado"*, que es bastante común en muchos textos recientes, pierde el sabor de las matemáticas, diluye su esencia y en lugar de ayudar a los estudiantes a descubrir y desarrollar sus habilidades matemáticas puede convertirse en una serie sin fin de ejemplos y ejercicios aburridos. El propósito de los "Fundamentos de Métodos Matemáticos para Física e Ingeniería" es dar a los estudiantes bases sólidas en métodos matemáticos, despertar su mente y enseñarles cierta cultura de estudios y aplicaciones de métodos matemáticos en su trabajo profesional. Se piensa que al alcanzar las metas propuestas por el texto, el estudiante se sentirá mucho más cómodo y con más confianza en sí mismo al estudiar otras materias y, en caso de que surja la necesidad, los conocimientos adquiridos por medio de este libro le permitirán avanzar de forma autodidacta en los estudios de otros métodos matemáticos, apoyándose en libros de texto más completos y avanzados.

Aprovechando el hecho de que ya hay muchos libros que contienen una gran cantidad de problemas de diferentes niveles sobre todos los temas que contiene este libro, los problemas y ejercicios de éste no son numerosos, sino escogidos

para ser altamente instructivos; no son únicamente para ejercitar al estudiante, sino para enriquecer y ampliar el conocimiento conceptual sobre el tema correspondiente.

Este libro fue pensado como un repaso conciso, autoconsistente y autosuficiente, una consulta para los que ya saben bastante sobre los métodos matemáticos y/o un curso intensivo (de uno o dos semestres) para los que quieren o necesitan elevar de manera adecuada el nivel de sus conocimientos en la materia.

La bibliografía es muy escasa e incluye básicamente los libros que se usaron para dar forma y contenido al texto.

Y, por fin, a los lectores les pido el gran favor de apuntar pero no contar mis errores (¡no los tiene aquel que no hace nada!). Tal vez mi error principal fue empezar el trabajo de escribir este libro, distrayéndome de la investigación científica que indudablemente regala muchos momentos de satisfacción. Justamente valorando el esfuerzo de la *tribu* llamada estudiantes sentí la necesidad de apoyar sus estudios con mi experiencia personal. El resultado está aquí y a ustedes se les otorga el pleno derecho de juzgarlo.

Se les agradece a Raymundo Mendoza Arce la elaboración de la portada y a Víctor Ayala Ramírez su asesoría en cuanto a la edición del texto en LaTeX.

**Dedicatoria**. A mi familia y a mis estudiantes, que me dan la razón para vivir una vida completa y feliz.

E.V. Kurmyshev

# Capítulo 1

# Ecuaciones diferenciales ordinarias

## 1.1. Clasificación y origen de las ecuaciones diferenciales

Una **ecuación diferencial** es una ecuación que contiene derivadas. La ecuación de la forma

$$F\left(x, y, y', y'', \ldots, y^{(n)}\right) = 0 \tag{1.1}$$

en donde $y = y(x)$ es función de la variable independiente $x$ que se busca y $y^{(n)}$ es la $n$-ésima derivada con respecto a $x$, se llama **ecuación diferencial ordinaria de orden $n$-ésimo**. El **orden** de una ecuación diferencial corresponde al orden de la derivada mayor que interviene en la ecuación, mientras que el **grado** de la ecuación diferencial denota el grado del exponente de la derivada de mayor orden (se implica que la ecuación está escrita en forma polinomial con respecto a la derivada mayor). Presentamos ejemplos de ecuaciones diferenciales:

$$y' = 5 + x \tag{1.2}$$

$$(y'')^2 + (y')^3 + 3y = x^2 \tag{1.3}$$

$$y'' + 3y' + 2y = 0 \tag{1.4}$$

$$y''' + 2(y'')^2 + y' = \cos x \tag{1.5}$$

$$z_{xx} + z_{yy} = x^2 + y \tag{1.6}$$

en donde $y' = dy/dx$ es la derivada ordinaria de la función $y(x)$ con respecto a la variable independiente $x$, $z_{xx} = \partial^2 z/\partial x^2$ es la derivada parcial de segundo orden. Las ecuaciones (1.2), (1.3), (1.4) y (1.5) son diferenciales ordinarias (EDO), mientras que la ecuación (1.6) es un ejemplo de ecuación diferencial parcial (EDP). La ecuación (1.2) es de orden 1 y grado 1, EDO (1,1); la ecuación (1.3) es de orden 2 y grado 2, EDO (2,2); la ecuación (1.4) es de orden 2 y grado

1, EDO (2,1); la ecuación (1.5) es de orden 3 y grado 1, EDO (3,1); la ecuación (1.6) es EDP (2,1).

Las EDO pueden surgir a partir de

a) problemas geométricos,

b) problemas de ciencias naturales,

c) primitivas de EDO.

Cualquier función $y = y(x)$ que, al sustituirla en la ecuación (1.1), transforma la ecuación diferencial en identidad, recibe el nombre de solución de esta ecuación. A menudo, la solución se da en forma implícita, $\Phi(x, y) = 0$. Una **primitiva** es una relación funcional entre variables $x$ y $y$ que contiene $n$ constantes arbitrarias $C_1, C_2, \ldots, C_n$:

$$\Phi(x, y, C_1, C_2, \ldots, C_n) = 0 \tag{1.7}$$

En general, una primitiva que contenga $n$ **constantes arbitrarias esenciales** (que no se pueden sustituir por un número menor de constantes) se puede reducir a una EDO de orden $n$, excluyendo las constantes $C_1, C_2, \ldots, C_n$ del sistema de ecuaciones obtenidas derivando la primitiva $n$ veces con respecto a la variable independiente:

$$\Phi(x, y, C_1, C_2, \ldots, C_n) = 0, \quad \frac{d\Phi}{dx} = 0, \quad \ldots, \quad \frac{d^n\Phi}{dx^n} = 0 \tag{1.8}$$

**Ejemplo 1.1.** *Derivando la primitiva $y = x^2 + C$ con respecto a $x$, excluimos la constante $C$ y obtenemos la EDO correspondiente, $y' = 2x$.*

**Ejemplo 1.2.** *La primitiva $y = C_1x^2 + C_2$ contiene dos constantes esenciales. Derivando una vez se excluye la constante $C_2$ y se tiene $y' = 2C_1x$; derivando nuevamente, se tiene $y'' = 2C_1$. Resolviendo esta ecuación para la constante $C_1$ y sustituyendo en la ecuación de la primera derivación, resulta la EDO $y' = y''x$ de segundo orden.*

## 1.2. Soluciones de ecuaciones diferenciales ordinarias

Resolver una EDO de orden $n$ implica encontrar una relación entre variables $x$ y $y$ que contenga $n$ constantes arbitrarias independientes y que satisfaga la EDO, es decir, obtener una primitiva. Por ejemplo, la ecuación $y''' = 0$ tiene como su solución o primitiva la función $y = Ax^2 + Bx + C$. La primitiva de una ecuación diferencial se denomina normalmente la **solución general** de la ecuación. Una **solución particular** de una EDO se obtiene de la primitiva dando valores determinados a las constantes arbitrarias. Así, soluciones particulares de la ecuación anterior son: $y = 0$ para $A = B = C = 0$; $y = 2x + 1$ para $A = 0$, $B = 2$, $C = 1$; $y = 3x^2$ para $A = 3$, $B = C = 0$. Geométricamente, la primitiva es la ecuación de una familia de curvas y una solución particular es la ecuación de una de las curvas.

Por regla, la solución de una EDO de orden $n$ está sujeta a **condiciones iniciales**, que son valores de la función dependiente y de sus $(n-1)$ primeras derivadas en un punto dado $x_0$: $y(x_0) = a_0$, $y'(x_0) = a_1$, $y''(x_0) = a_2$, ..., $y^{(n-1)}(x_0) = a_{n-1}$, en donde $a_0, \ldots, a_{n-1}$ son constantes. Al resolver una EDO con condiciones iniciales, se encuentra uno ante los **problemas de existencia y unicidad** de la solución, es decir: ¿bajo qué condiciones se tiene al menos una solución? Y, ¿bajo qué condiciones la EDO tiene sólo una solución? En el caso de una EDO (1,1), la respuesta se da (sin demostración) por el siguiente teorema de existencia y unicidad.

**Teorema 1.1.** *Dada la EDO de primer orden y primer grado: $y' = g(x,y)$. Si $g(x,y)$ es continua en un dominio $R = \{(x,y)\}$ y $\partial g/\partial y$ existe y es continua en $R$, entonces existe una primitiva (una familia infinita de soluciones), $f(x,y,C) = 0$, siendo $C$ una constante arbitraria, tal que $y' = g(x,y)$. Además, la EDO con condiciones iniciales tiene sólo una solución tal que para cualquier punto $(x_0,y_0)$ se satisfacen las condiciones iniciales: $y(x_0) = y_0$, $y'(x_0) = g(x_0,y_0)$.*

Una EDO (1,1), $y' = g(x,y)$, asocia con cada punto $(x_0,y_0)$ de la región $R = \{(x,y)\}$ una dirección $m = y'(x_0,y_0) = g(x_0,y_0)$. La dirección en cada punto es la de la tangente a la curva de la familia, $f(x,y,C) = 0$, que pasa por este punto. La región $R$ con la dirección en cada uno de sus puntos se llama **campo de direcciones** de la ecuación diferencial $y' = g(x,y)$ y se representa generalmente por medio de un sistema de flechas con un ángulo de inclinación $\tan\alpha = g(x,y)$. Las curvas $K = g(x,y)$, en cuyos puntos la inclinación del campo tiene un valor constante $K$, se llaman **isoclinas**.

## 1.3. Ecuaciones diferenciales de primer orden y primer grado

Una EDO (1,1) se puede escribir en una de las dos formas:

$$M(x,y)dx + N(x,y)dy = 0 \quad (1.9)$$

$$y' = g(x,y) \quad (1.10)$$

**Ejemplo 1.3.** *La ecuación $y' = 5+x$ se puede escribir así: $(5+x)\,dx - dy = 0$, donde $M = 5+x$, $N = -1$. La ecuación $y' = (y+x)\,/\,(y-x)$ se puede escribir así: $(y+x)\,dx + (x-y)\,dy = 0$, donde $M = y+x$, $N = x-y$.*

Si $M(x,y)dx + N(x,y)dy$ es una diferencial completa de una función $\mu(x,y)$, de modo que $d\mu(x,y) = Mdx + Ndy$, la ecuación (1.9) se llama una **ecuación diferencial exacta** y $\mu(x,y) = C$ es su primitiva o solución general.

**Ejemplo 1.4.** *Sea $\mu(x,y) = x^3y^2$, entonces, la diferencial completa de esta función es $d\mu(x,y) = 3x^2y^2dx + 2x^3ydy$. La ecuación diferencial exacta es $3x^2y^2dx + 2x^3ydy = 0$ y, por lo tanto, la solución general de ésta es $x^3y^2 = C$.*

*Nótese que la ecuación diferencial a su vez se puede escribir como* $y' = \frac{-3x^2y^2}{2x^3y} = -\frac{3y}{2x}$.

La diferencial completa de una función $\mu(x,y)$ es

$$d\mu(x,y) = \frac{\partial\mu}{\partial x}dx + \frac{\partial\mu}{\partial y}dy \tag{1.11}$$

y para las segundas derivadas parciales se cumple la igualdad $\frac{\partial^2\mu}{\partial y\partial x} = \frac{\partial^2\mu}{\partial x\partial y}$. Por lo tanto, el criterio de que la forma diferencial $M(x,y)dx + N(x,y)dy$ es una diferencial completa se da por la igualdad

$$\frac{\partial M(x,y)}{\partial y} = \frac{\partial N(x,y)}{\partial x} \tag{1.12}$$

Si la ecuación (1.9) no es exacta, sin embargo se encuentra una función $\xi(x,y)$ tal que

$$\xi(x,y)(Mdx + Ndy) = d\mu(x,y) \tag{1.13}$$

la función $\xi(x,y)$ se llama un **factor integrante** y $\mu(x,y) = C$ es la primitiva de la ecuación (1.9).

**Ejemplo 1.5.** *La ecuación* $3ydx + 2xdy = 0$ *no es exacta, pero al multiplicarla por* $x^2y$*, se tiene la ecuación* $3x^2y^2dx + 2x^3ydy = 0$*, que es exacta con la primitiva (su solución general)* $x^3y^2 = C$.

Si la ecuación (1.9) no es exacta y no se encuentra rápidamente un factor integrante, es posible que mediante un cambio de una o de las dos variables se obtenga una ecuación en la que se pueda hallar un factor integrante. Aunque no se puede dar una regla general para hallar un factor integrante o una transformación, se pueden considerar tres agrupaciones importantes, que son en cierto modo convencionales, para obtener las soluciones:

1) Ecuaciones del tipo de separación de variables y reducibles a ellas,

2) ecuaciones diferenciales exactas y reducibles a ellas,

3) EDO (1,1) lineales, de la forma $y' + P(x)y = Q(x)$.

**Ejemplo 1.6.** *La ecuación diferencial* $xdy - ydx = 0$ *se puede clasificar en cualquiera de los grupos: 1) las variables se separan multiplicando la ecuación por el factor integrante* $1/xy$*, así que* $dy/y - dx/x = 0$*, de donde* $\ln y - \ln x = \ln C$*, o sea,* $y/x = C$*; 2) por medio del factor integrante* $1/x^2$*, la ecuación se convierte en exacta,* $(xdy - ydx)/x^2 = 0$*, de donde la primitiva es* $y/x = C$*; por medio del factor integrante* $1/y^2$*, la ecuación se convierte en exacta,* $(xdy - ydx)/y^2 = 0$ *con la primitiva* $-x/y = C_1$ *o* $y/x = -1/C_1 = C$*; 3) escrita en la forma* $y' - y/x = 0$*, es la ecuación lineal de primer orden.*

Llama la atención el hecho de que la forma de la primitiva no es única. La primitiva en el ejemplo anterior puede ponerse en las siguientes formas: a) $\ln y - \ln x = \ln C$, b) $y/x = C$, c) $y = Cx$, d) $x/y = K$. Aceptando cualquiera de estas formas, se pueden perder ciertas soluciones particulares.

### 1.3.1. Separación de variables

Las EDO (1,1) de la forma

$$y' = f(x)\,g(y) \quad \text{o} \quad X(x)dx + Y(y)dy = 0 \tag{1.14}$$

se llaman ecuaciones diferenciales con **variables separables**. Dividiendo la primera ecuación por $g(y)$ y multiplicando por $dx$, obtenemos $dy/g(y) = f(x)\,dx$. Integrando la ecuación (1.14), se obtiene la solución general en la forma

$$\int \frac{dy}{g(y)} = \int f(x)\,dx + C \quad \text{o} \quad \int X(x)dx + \int Y(y)dy = C \tag{1.15}$$

Se halla rápidamente un factor integrante para la ecuación de la forma

$$f_1(x)g_2(y)dx + f_2(x)g_1(y)dy = 0 \tag{1.16}$$

Al multiplicarla por el factor integrante $\xi(x,y) = \frac{1}{g_2(y)f_2(x)}$, la ecuación resulta en la forma con variables separadas

$$\frac{f_1(x)}{f_2(x)}dx + \frac{g_1(y)}{g_2(y)}dy = 0 \tag{1.17}$$

cuya primitiva se obtiene simplemente integrando:

$$\int \frac{f_1(x)}{f_2(x)}dx + \int \frac{g_1(y)}{g_2(y)}dy = C \tag{1.18}$$

**Ejemplo 1.7.** *La ecuación diferencial* $(x-1)^2ydx + x^2(y+1)dy = 0$ *que no es exacta, multiplicada por el factor integrante* $\frac{1}{yx^2}$ *resulta en* $\frac{(x-1)^2}{x^2}dx + \frac{(y+1)}{y}dy = 0$. *Integrando se obtiene la solución general,* $x - 2\ln x - \frac{1}{x} + y + \ln y = C$.

### 1.3.2. Ecuaciones diferenciales homogéneas y reducción a separación de variables

Una función $f(x,y)$ es homogénea de grado $n$ si para un número entero $n$ se cumple la igualdad

$$f(\lambda x, \lambda y) = \lambda^n f(x,y) \tag{1.19}$$

**Ejemplo 1.8.** *La función* $f(x,y) = x^3 - x^2y$ *es de tercer grado. La* $f(x,y) = \exp(y/x) + \tan(x/y)$ *es de grado 0. La función* $f(x,y) = x^3 - x^2 \operatorname{sen} y$ *no es homogénea.*

Una EDO (1,1), $M(x,y)dx + N(x,y)dy = 0$, es una **ecuación diferencial homogénea** si tanto $M(x,y)$ como $N(x,y)$ son funciones homogéneas del mismo grado $n$.

**Ejemplo 1.9.** *La EDO (1,1)* $x\ln(y/x)\,dx + (y^2/x)\operatorname{arcsen}(y/x)\,dy = 0$ *es la ecuación diferencial homogénea, dado que tiene funciones* $M(x,y)$ *y* $N(x,y)$ *homogéneas de primer grado.*

**Proposición 1.1.** *La transformación*

$$y = vx, \ \ con \ dy = xdv + vdx \tag{1.20}$$

*reduce cualquier EDO (1,1) homogénea, $M(x,y)dx + N(x,y)dy = 0$, a una ecuación del tipo de variables separables.*

**Demostración.** *Dado que $M(x,y)$ y $N(x,y)$ son homogéneas de grado n y considerando $\lambda = \frac{1}{x}$, tenemos*

$$M(\lambda x, \lambda y) = \lambda^n M(x,y) \quad \Rightarrow \quad M(1, \frac{y}{x}) = \frac{1}{x^n} M(x,y)$$

$$N(\lambda x, \lambda y) = \lambda^n N(x,y) \quad \Rightarrow N(1, \frac{y}{x}) = \frac{1}{x^n} N(x,y)$$

*Empleando la notación $N_1(\frac{y}{x}) \equiv N(1, \frac{y}{x})$, $M_1(\frac{y}{x}) \equiv M(1, \frac{y}{x})$, reescribimos la ecuación inicial como*

$$x^n \left[ M_1\left(\frac{y}{x}\right) dx + N_1\left(\frac{y}{x}\right) dy \right] = 0 \quad \Rightarrow M_1\left(\frac{y}{x}\right) dx + N_1\left(\frac{y}{x}\right) dy = 0$$

*Aplicando la transformación $y = vx$, $dy = xdv + vdx$ a la ecuación anterior se tiene*

$$[M_1(v) + vN_1(v)]\, dx + N_1(v) x dv = 0$$

*que tiene la forma*

$$f(v)dx + xN_1(v)dv = 0 \tag{1.21}$$

*Al multiplicar por el factor integrante*

$$\frac{1}{f(v)x} = \frac{1}{x\,[M_1(v) + vN_1(v)]} \tag{1.22}$$

*resulta en la ecuación*

$$\frac{dx}{x} + \frac{N_1(v)}{M_1(v) + vN_1(v)} dv = 0 \tag{1.23}$$

*cuya solución general se obtiene integrando,*

$$\int \frac{dx}{x} + \int \frac{N_1(v)}{M_1(v) + vN_1(v)} dv = C \tag{1.24}$$

*Quedando completa la demostración.* ■

**Problema.** *Sean $f_1(x,y)$ y $f_2(x,y)$ dos funciones homogéneas de grado k y m, respectivamente. ¿Qué se puede decir sobre la homogeneidad de las funciones $g(x,y) = f_1(x,y) f_2(x,y)$ y $h(x,y) = \alpha f_1(x,y) + \beta f_2(x,y)$, en donde $\alpha$ y $\beta$ son constantes?*

**Problema.** *Sea $f(x,y)$ una función homogénea de grado n. Demuestre la igualdad (el teorema de Euler),*

$$xf_x + yf_y = nf(x,y)$$

*en donde $f_x$ y $f_y$ son derivadas parciales.*

### 1.3.3. Ecuaciones con $M(x,y)$ y $N(x,y)$ lineales, pero no homogéneas

La ecuación no homogénea

$$(a_1x + b_1y + c_1)dx + (a_2x + b_2y + c_2)dy = 0 \qquad (1.25)$$

se reduce por una transformación lineal de variables a la forma en la que las variables son separables. Las funciones $g_1(x,y) \equiv a_1x+b_1y+c_1 = 0$ y $g_2(x,y) \equiv a_2x+b_2y+c_2 = 0$ definen dos rectas que en el caso de que $c_1 = c_2 = 0$ pasan por el origen del sistema de coordenadas. Notando que con $c_1 = c_2 = 0$ la ecuación (1.25) será homogénea de grado uno, encontramos el punto de intersección $(h,k)$ de las rectas $g_1(x,y) = 0$ y $g_2(x,y) = 0$ y movemos allí el origen del nuevo sistema de coordenadas, quedando la transformación lineal $x = X+h$, $y = Y+k$. En las nuevas coordenadas las ecuaciones de dichas rectas son

$$a_1(X+h) + b_1(Y+k) + c_1 = a_1X + b_1Y + a_1h + b_1k + c_1 = 0$$

$$a_2(X+h) + b_2(Y+k) + c_2 = a_2X + b_2Y + a_2h + b_2k + c_2 = 0$$

y el punto de intersección se define a condición de que $X = Y = 0$, es decir, que $a_1h + b_1k + c_1 = 0$ y $a_2h + b_2k + c_2 = 0$. Dado que $a_1b_2 - a_2b_1 \neq 0$, de las últimas dos ecuaciones se obtienen unívocamente $(h,k)$. Por la transformación lineal $x = X+h$, $y = Y+k$, la ecuación (1.25) se reduce a la ecuación homogénea de grado uno

$$(a_1X + b_1Y)\,dX + (a_2X + b_2Y)\,dY = 0 \qquad (1.26)$$

que a su vez se reduce a una ecuación del tipo de separación de variables.

Cuando $a_1b_2 - a_2b_1 = 0$, la ecuación (1.25) se reduce por la transformación

$$a_1x + b_1y = t, \quad dy = (dt - a_1dx)/b_1$$

a la forma $P(t)\,dx + Q(t)\,dt = 0$ en la que las variables son separables.

### 1.3.4. Ecuaciones de la forma $f(xy)ydx + g(xy)xdy = 0$

La transformación $xy = z$, reduce una ecuación de este tipo a la forma de separación de variables. Sea $xy = z$. Sustituyendo $y = \frac{z}{x}$ y $dy = \frac{xdz - zdx}{x^2}$ en la ecuación original, se tiene

$$\left(\frac{z}{x}\right) f(z)dx + xg(z)\left(\frac{xdz - zdx}{x^2}\right) = 0 \qquad (1.27)$$

de donde se obtiene la ecuación en la que las variables son separables:

$$z\,(f(z) - g(z))\,dx + xg(z)dz = 0 \Rightarrow \frac{1}{x}dx + \frac{g(z)}{z\,(f(z) - g(z))}dz = 0 \qquad (1.28)$$

### 1.3.5. Ecuaciones diferenciales exactas y reducción a ellas

Una ecuación diferencial de la forma de la ecuación (1.9),

$$M(x,y)dx + N(x,y)dy = 0$$

se llama exacta si existe una función $\mu(x,y)$ tal que

$$d\mu(x,y) = Mdx + Ndy \tag{1.29}$$

en cuyo caso $\mu(x,y) = C$ es la primitiva de la ecuación.

La condición suficiente y necesaria de que la forma diferencial $Mdx + Ndy$ sea la diferencial completa de una función $\mu(x,y)$ se da por la ecuación (1.12),

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$$

En este caso,

$$M(x,y) = \frac{\partial \mu}{\partial x}, \quad N(x,y) = \frac{\partial \mu}{\partial y} \tag{1.30}$$

**Ejemplo 1.10.** *Sea $(x^2 - y)dx + (-x + y^2)dy = 0$, entonces $\frac{\partial(x^2-y)}{\partial y} = -1$, y $\frac{\partial(-x+y^2)}{\partial x} = -1$, por lo que la ecuación es exacta.*

La integral $\int_A^B d\mu(x,y) = \mu(x,y) - \mu(x_a, y_a)$ no depende de la trayectoria, donde $A(x_a, y_a)$ y $B(x,y)$ son dos puntos cualesquiera en el plano $xy$. Por lo tanto, integrando con respecto a $x$ y siendo $y$ constante, tenemos

$$\mu(x,y) = \int^x \frac{\partial \mu}{\partial x} dx + \varphi(y) = \int^x M(x,y)\,dx + \varphi(y) \tag{1.31}$$

donde la función $\varphi(y)$ es la "constante" de integración con respecto a $x$ y que aún no está determinada. Luego,

$$\frac{\partial \mu}{\partial y} = \frac{\partial}{\partial y}\left[\int^x M(x,y)dx + \varphi(y)\right] = \frac{\partial}{\partial y}\left[\int^x M(x,y)dx\right] + \varphi'(y) = N(x,y) \tag{1.32}$$

De la última ecuación podemos encontrar $\varphi(y)$, integrando con respecto a $y$ y tratando $x$ como una constante:

$$\varphi(y) = \int^y N(x,y)dy - \int^y \frac{\partial}{\partial y}\left[\int^x M(x,y)dx\right] dy$$

Sustituyendo la $\varphi(y)$ en la ecuación (1.31), encontraremos la primitiva de la ecuación diferencial exacta:

$$\mu(x,y) = \int^x M(x,y)\,dx + \int^y N(x,y)dy - \int^y \frac{\partial}{\partial y}\left[\int^x M(x,y)dx\right] dy = C \tag{1.33}$$

**Ejemplo 1.11.** *La ecuación* $(2x^3+3y)dx+(3x+y+1)dy=0$ *es exacta, pues* $\frac{\partial M}{\partial y}=\frac{\partial N}{\partial x}=3$. *Luego,*

$$M(x,y)=\frac{\partial \mu}{\partial x}=2x^3+3y, \qquad N(x,y)=\frac{\partial \mu}{\partial y}=3x+y+1$$

*así que*

$$\mu(x,y)=\int^x (2x^3+3y)\,dx+\varphi(y)=\frac{x^4}{2}+3xy+\varphi(y)$$

$$N(x,y)=\frac{\partial \mu}{\partial y}=3x+\varphi'(y)=3x+y+1$$

*de donde* $\varphi'(y)=y+1$, *que integrando resulta en* $\varphi(y)=\frac{y^2}{2}+y$. *Finalmente, la primitiva de la ecuación es*

$$\mu(x,y)=\frac{x^4}{2}+3xy+\frac{y^2}{2}+y=C$$

Si la ecuación (1.9) no es exacta, se busca algún factor integrante.

**Proposición 1.2.** *Si* $\frac{M_y-N_x}{N}=f(x)$, *función sólo de* $x$, *entonces* $\exp\left[\int f(x)dx\right]$ *es un factor integrante de la ecuación (1.9).*

**Demostración.** *Verificamos la condición necesaria y suficiente. El factor* $\xi(x)=\exp\left[\int f(x)dx\right]$ *es una función sólo de* $x$. *Entonces,*

$$\frac{\partial}{\partial y}\left(\xi(x)\,M(x,y)\right)=\xi M_y$$

$$\frac{\partial}{\partial x}\left(\xi(x)\,N(x,y)\right)=\xi N_x+\xi f N=\xi N_x+\xi\left(M_y-N_x\right)=\xi M_y$$

*Esto completa la demostración.* ■

**Proposición 1.3.** *Si* $\frac{M_y-N_x}{M}=-g(y)$, *función sólo de* $y$, *entonces* $\exp\left[\int g(y)dy\right]$ *es un factor integrante de la ecuación (1.9).*

**Demostración.** *El factor* $\xi(y)=\exp\left[\int g(y)dy\right]$ *es función sólo de* $y$. *Al multiplicar por el factor integrante,* $d\mu(x,y)=\xi\,Mdx+\xi\,Ndy$, *se tiene que cumplir la condición de que* $\frac{\partial(\xi\ M)}{\partial y}=\frac{\partial(\xi\ N)}{\partial x}$. *Diferenciando,*

$$\frac{\partial(\xi\ M)}{\partial y}=\xi M_y+\xi g M=\xi M_y+\xi\left(N_x-M_y\right)=\xi N_x$$

$$\frac{\partial(\xi\ N)}{\partial x}=\xi N_x$$

*Por lo tanto, queda demostrado que el factor es integrante.* ■

**Proposición 1.4.** *Si la ecuación (1.9) es homogénea y* $Mx+Ny \neq 0$, *entonces,* $(Mx+Ny)^{-1}$ *es un factor integrante.*

**Demostración.** *Sea* $\xi(x,y) = (Mx+Ny)^{-1}$. *Se tiene*

$$\frac{\partial(\xi\ M)}{\partial y} = \frac{\partial}{\partial y}\frac{M}{Mx+Ny} = \frac{yM_yN - yN_yM - MN}{(Mx+Ny)^2}$$

$$\frac{\partial(\xi\ N)}{\partial x} = \frac{\partial}{\partial x}\frac{N}{Mx+Ny} = \frac{xN_xM - xM_xN - MN}{(Mx+Ny)^2}$$

*Entonces, usando el teorema de Euler de una función* $H(x,y)$ *homogénea de grado* $n$:

$$xH_x + yH_y = nH$$

*se tiene*

$$\frac{\partial(\xi\ M)}{\partial y} - \frac{\partial(\xi\ N)}{\partial x} = \frac{(yM_y + xM_x)N - (yN_y + xN_x)M}{(Mx+Ny)^2} = 0$$

*y queda demostrado que el factor es integrante.* ∎

Nótese que una EDO (1,1) homogénea siempre puede ser resuelta reduciéndola a la ecuación del tipo de separación de variables por medio de la transformación de variables, ecuación (1.20). Además, puesto que $M\ x + N\ y \neq 0$, ésta se resuelve reduciéndola a la exacta por medio de un factor integrante.

A veces un factor integrante se halla por inspección, agrupando convenientemente los términos de la ecuación y reconociendo un cierto grupo de términos como parte de una diferencial exacta. En la siguiente tabla se presentan algunos factores integrantes de utilidad.

| **Grupo de términos** | **Factor integrante** | **Primitiva de ecuación** |
|---|---|---|
| $-ydx + xdy$ | $\frac{1}{x^2}$ | $\frac{y}{x} = C$ |
| $-ydx + xdy$ | $\frac{1}{y^2}$ | $-\frac{x}{y} = C$ |
| $-ydx + xdy$ | $\frac{1}{xy}$ | $\ln\frac{y}{x} = C$ |
| $-ydx + xdy$ | $\frac{1}{x^2+y^2}$ | $\arctan\frac{y}{x} = C$ |
| $ydx + xdy$ | $\frac{1}{xy}$ | $\ln xy = C$ |
| $xdx + ydy$ | $\frac{1}{x^2+y^2}$ | $\frac{1}{2}\ln\left(x^2+y^2\right) = C$ |

### 1.3.6. Ecuaciones diferenciales lineales y reducibles a lineales

La ecuación

$$y' + P(x)y = Q(x) \tag{1.34}$$

se llama **ecuación diferencial lineal de primer orden**, siendo una forma lineal tanto en la variable dependiente como en su derivada.

**Ejemplo 1.12.** *La ecuación* $y' + 3x^2y = \operatorname{sen} x$, *es lineal, mientras que* $y' + 3x^2y^2 = \operatorname{sen} x$, *no lo es.*

La ecuación (1.34) escrita en diferenciales es $(yP - Q)dx + dy = 0$. Luego, $M(x,y) = yP(x) - Q(x)$ y $N(x,y) = 1$ resultando en que $\frac{M_y - N_x}{N} = P(x)$ es sólo función de $x$. Por lo tanto, según la Proposición 18, un factor integrante de la ecuación (1.34) es $\xi(x) = e^{\int P(x)dx}$. La ecuación diferencial, multiplicada por el factor integrante, resulta en

$$e^{\int P(x)dx}(y' + yP(x)) = \frac{d}{dx}\left[ye^{\int P(x)dx}\right] = Q(x)e^{\int P(x)dx}$$

y la primitiva de la ecuación (1.34) es

$$ye^{\int P(x)dx} = \int Q(x)e^{\int P(x)dx}dx + C \tag{1.35}$$

### 1.3.7. Ecuación de Bernoulli

La ecuación diferencial

$$y' + P(x)y = Q(x)y^a \tag{1.36}$$

donde $a$ es cualquier número, se llama **ecuación de Bernoulli**. Para $a = 0$ y $a = 1$, la ecuación es lineal; de lo contrario no. Una forma equivalente es $y^{-a}y' + y^{1-a}P(x) = Q(x)$. Haciendo el cambio de variable

$$z = y^{1-a} \tag{1.37}$$

se tiene que $z' = (y^{1-a})' = (1-a)y^{-a}y'$. Al sustituir en la ecuación (1.36), se tiene

$$z' + (1-a)zP(x) = (1-a)Q(x) \tag{1.38}$$

que es una ecuación lineal.

El factor integrante $e^{(1-a)\int P(x)dx}$, aplicado a la ecuación (1.38), lleva a la solución

$$z = y^{1-a} = e^{(a-1)\int P(x)dx}\left[\int (1-a)Q(x)e^{(1-a)\int P(x)dx}dx\right] + Ce^{(a-1)\int P(x)dx} \tag{1.39}$$

### 1.3.8. Ecuación de Riccati

La ecuación

$$y' + P(x)y + Q(x)y^2 + R(x) = 0 \tag{1.40}$$

donde $P(x)$, $Q(x)$ y $R(x)$ son funciones de $x$, es conocida como **ecuación de Riccati generalizada**. La ecuación de Riccati original es $y' + Ay^2 = Bx^m$, donde $A$ y $B$ son constantes y $m$ es un número natural. En caso de $R(x) = 0$, la ecuación (1.40) es de Bernoulli con $a = 2$. Si $Q(x) = 0$, se tiene una ecuación lineal.

En general, la ecuación de Riccati generalizada no es integrable en cuadraturas. Sin embargo, cuando se conoce cualquier solución particular, la solución

general de ésta se obtiene fácilmente reduciéndola a la ecuación de Bernoulli, y esta última es reducida a una ecuación lineal.

Sea $y_1$ una solución particular de (1.40). Buscamos la solución general de la ecuación (1.40) en la forma

$$y = y_1 + z \tag{1.41}$$

Sustituyendo (1.41) en (1.40), se tiene

$$z' + Q(x)\left[2y_1 z + z^2\right] + P(x)z = 0 \quad \text{o} \quad z' + \tilde{P}(x)z = \tilde{Q}(x)z^2 \tag{1.42}$$

que es la ecuación de Bernoulli, ecuación (1.36), con respecto a la incógnita $z$, donde $\tilde{Q}(x) = -Q(x)$, $\tilde{P}(x) = P(x) + 2y_1 Q(x)$ y $a = 2$. El cambio de variable,

$$v = \frac{1}{z}$$

reduce la última ecuación a una ecuación lineal (véase la sección anterior).

**Proposición 1.5.** *Sean $y_1$, $y_2$, $y_3$ soluciones particulares de (1.40) y sea $y$ su solución general. Entonces,*

$$u = \frac{1}{y - y_1},\; u_1 = \frac{1}{y_2 - y_1},\; u_2 = \frac{1}{y_3 - y_1}$$

*son soluciones de la misma ecuación diferencial lineal y debido a esto*

$$\frac{u - u_1}{u_2 - u_1} = C \quad \Rightarrow \quad \frac{y - y_2}{y - y_1} = C\,\frac{y_3 - y_2}{y_3 - y_1}$$

*donde $C$ es una constante.*

La última fórmula muestra que la solución general de la ecuación de Riccati se puede expresar de la manera racional en términos de cualquiera de tres soluciones particulares distintas; además, se tiene que la solución general es una función racional de la constante de integración. Por tanto, cualquier función del tipo

$$y = \frac{Cf_1 + f_2}{Cf_3 + f_4} \tag{1.43}$$

satisface una ecuación de Riccati, ecuación (1.40), donde $C$ es una constante arbitraria y $f_1$, $f_2$, $f_3$ y $f_4$ son funciones dadas de $x$. Lo último se demuestra eliminando $C$ de las expresiones para $y$ y su derivada $y'$.

Con $Q(x) = 0$ la ecuación de Riccati se transforma en una ecuación lineal. Si $Q(x) \neq 0$, la ecuación de Riccati es reducible a una EDO (2,1) lineal. Haciendo el cambio de variable dependiente,

$$y = \frac{v}{Q(x)} \tag{1.44}$$

en la ecuación (1.40), resulta

$$v' + v^2 + v\left[P - \frac{Q'}{Q}\right] + R\,Q = 0$$

Luego, sustituyendo $v = (\ln u)' = \frac{u'}{u}$, se tiene

$$\frac{u''u - u'^2}{u^2} + \left(\frac{u'}{u}\right)^2 + \frac{u'}{u}\left[P - \frac{Q'}{Q}\right] + R\,Q = 0$$

y finalmente

$$u'' + u'\left[P - \frac{Q'}{Q}\right] + uR\,Q = 0 \tag{1.45}$$

que es la ecuación diferencial lineal de segundo orden.

**Ejemplo 1.13.** *La ecuación $y' - y^2 + 2xy - \left(x^2 + 1\right) = 0$ es de Riccati. Es fácil ver que $y = x$ es una solución particular. Haciendo la sustitución $y = x + z$, se tiene*

$$1 + z' - \left(x^2 + 2xz + z^2\right) + 2x(x + z) - (x^2 + 1) = 0$$

$$z' = z^2$$

*La última ecuación se resuelve fácilmente por separación de variables,*

$$\frac{dz}{dx} = z^2 \quad \Rightarrow \quad \int \frac{dz}{z^2} = \int dx \quad \Rightarrow \quad x = -\frac{1}{z} + C$$

*También puede resolverse como ecuación de Bernoulli realizando la sustitución $v = \frac{1}{z}$, $z = \frac{1}{v}$, $z' = -\frac{v'}{v^2}$,*

$$-\frac{v'}{v^2} = \frac{1}{v^2} \quad \Rightarrow \quad v' = -1 \quad \Rightarrow \quad v = -x + C \quad \Rightarrow \quad -x + C = \frac{1}{z}$$

*Así, $z = \frac{1}{C - x}$, luego, como $y = x + z$, se tiene finalmente la solución general*

$$y = \frac{Cx - x^2 + 1}{C - x}$$

**Ejemplo 1.14.** *La ecuación $y' - y^2 + \frac{y}{x} + \frac{1}{x^2} = 0$ es de Riccati. Se puede comprobar fácilmente que $y_1 = \frac{1}{x}$ es una solución particular. Sustituyendo $y = y_1 + z = \frac{1}{x} + z$ en la ecuación, se tiene*

$$\left(\frac{1}{x}\right)' + z' - \left(\frac{1}{x^2} + \frac{2z}{x} + z^2\right) + \frac{1}{x^2} + \frac{z}{x} + \frac{1}{x^2} = 0$$

*o bien, reacomodando los términos se tiene la ecuación de Bernoulli con $a = 2$:*

$$z' - \frac{z}{x} = z^2$$

*El cambio de la variable dependiente, $u = z^{-1}$, pues $z = u^{-1}$ y $z' = -u^{-2}u'$, reduce la ecuación anterior a la ecuación lineal*

$$u' + \frac{1}{x}u = -1$$

*cuyo factor integrante es*

$$\xi(x) = e^{\int \frac{1}{x}dx} = e^{\ln x} = x$$

*Al multiplicar la ecuación diferencial por el factor integrante se tiene*

$$xu' + u = -x, \quad \int d(xu) = -\int x dx, \quad xu = -\frac{x^2}{2} + C$$

*de donde* $u = -\frac{x}{2} + \frac{C}{x}$*. Ahora, como* $z = u^{-1}$*, tenemos que*

$$z = \frac{2x}{2C - x^2}$$

*Finalmente, como* $y = \frac{1}{x} + z$*, se tiene que la solución general es*

$$y = \frac{x^2 + 2C}{x(2C - x^2)}$$

## 1.4. Ecuaciones de primer orden y de grado superior

Una ecuación diferencial de primer orden tiene la forma $f(x, y, y') = 0$, o bien, $f(x, y, p) = 0$, donde sustituimos $y'$ por $p$, $y' = p$. Por ejemplo, la ecuación $p^2 - 3px + 2y = 0$, es una EDO (1,2). La ecuación diferencial ordinaria de primer orden de grado $n$ se puede escribir en la forma

$$p^n + P_1(x, y)p^{n-1} + P_2(x, y)p^{n-2} + \cdots + P_{n-1}(x, y)p + P_n(x, y) = 0 \qquad (1.46)$$

Al resolver estas ecuaciones, EDO (1,n), pueden presentarse varios casos.

### 1.4.1. Ecuación resuelta con respecto a $y'$

Supongamos que la ecuación (1.46), considerada como un polinomio en $p$, se puede resolver con respecto a $p$, es decir, esta ecuación polinomial se puede presentar en la forma de $n$ factores

$$(p - F_1(x, y))(p - F_2(x, y)) \quad \cdots \quad (p - F_n(x, y)) = 0 \qquad (1.47)$$

donde las raíces de la ecuación (1.46), $F_i(x, y)$, $i = 1, 2, \ldots, n$, son funciones de $x$ y $y$. Igualando a cero cada factor y resolviendo las $n$ ecuaciones diferenciales resultantes de primer orden y primer grado

$$\frac{dy}{dx} = F_1(x, y), \quad \frac{dy}{dx} = F_2(x, y), \quad \ldots, \quad \frac{dy}{dx} = F_n(x, y)$$

se obtienen las primitivas correspondientes

$$f_1(x, y, C) = 0, \quad f_2(x, y, C) = 0, \quad \ldots, \quad f_n(x, y, C) = 0 \qquad (1.48)$$

Cada una de estas primitivas anula uno de los factores en la ecuación (1.47) y, por esto satisface tanto la ecuación (1.47) como la ecuación (1.46). Por lo tanto, cada una de las primitivas (1.48) puede llamarse la primitiva parcial de las ecuaciones (1.47, 1.46). La primitiva de (1.47) y, por tanto, de (1.46), es el producto

$$f_1(x,y,C)f_2(x,y,C) \quad \cdots \quad f_n(x,y,C)=0 \tag{1.49}$$

de $n$ soluciones, ecuación (1.48).

**Ejemplo 1.15.** *Es fácil ver que la EDO*

$$p^2-2x\,(y+1)\,p+4x^2y=0, \quad \textit{o bien,} \quad (p-2x)\,(p-2xy)=0$$

*tiene las siguientes primitivas parciales*

$$\begin{aligned} y &= x^2+C, \quad \textit{o bien,} \quad f_1\,(x,y,C)\equiv y-x^2-C=0 \\ y &= C\exp\left(x^2\right), \quad \textit{o bien,} \quad f_2\,(x,y,C)\equiv y-C\exp\left(x^2\right)=0 \end{aligned}$$

*que son soluciones generales de las EDO (1,1), $p-2x=0$ y $p-2xy=0$, respectivamente. Entonces, la primitiva completa es*

$$f_1\,(x,y,C)\,f_2\,(x,y,C)\equiv\left(y-x^2-C\right)\left(y-C\exp\left(x^2\right)\right)=0$$

## 1.4.2. Ecuación resuelta con respecto a $y$

Supongamos que la EDO de primer orden se resuelve con respecto a $y$,

$$y=f(x,p) \tag{1.50}$$

Derivando la ecuación (1.50) con respecto a $x$, hallamos la ecuación que, con respecto a $p$ y $dp/dx$, es una EDO (1,1):

$$\frac{dy}{dx}=p=\frac{\partial f}{\partial x}+\frac{\partial f}{\partial p}\frac{dp}{dx} \quad \Rightarrow \quad \frac{dp}{dx}=F\,(x,p) \tag{1.51}$$

donde $F\,(x,p)=\left(p-\frac{\partial f}{\partial x}\right)/\frac{\partial f}{\partial p}$. Resuélvase la ecuación (1.51) obteniendo la primitiva $\phi(x,p,C)=0$. Después, obténgase la primitiva $\Phi(x,y,C)=0$ excluyendo $p$ entre las dos ecuaciones,

$$y=f(x,p) \quad \text{y} \quad \phi(x,p,C)=0$$

cuando sea posible, o bien se quedan $x$ y $y$ expresadas como funciones del parámetro $p$.

## 1.4.3. Ecuación resuelta con respecto a $x$

Cuando la ecuación (1.46) se resuelve con respecto a $x$,

$$x=f(y,p) \tag{1.52}$$

derivando la ecuación (1.52) con respecto a $y$, se tiene una EDO (1,1) con respecto a $p$ y $dp/dy$:

$$\frac{dx}{dy} = \frac{1}{p} = \frac{\partial f}{\partial y} + \frac{\partial f}{\partial p}\frac{dp}{dy} \Rightarrow \frac{dp}{dy} = F\left(y,p\right) \tag{1.53}$$

donde $F\left(x,p\right) = \left(\frac{1}{p} - \frac{\partial f}{\partial y}\right) / \frac{\partial f}{\partial p}$. Hallando la primitiva $\phi(y,p,C) = 0$ de la ecuación (1.53) y excluyendo $p$, cuando sea posible, de las dos ecuaciones,

$$x = f(y,p) \quad \text{y} \quad \phi(y,p,C) = 0$$

encontramos la primitiva $\Phi(x,y,C) = 0$.

**Ejemplo 1.16.** *Considérese la ecuación*

$$y = 3px + 6p^2y^2$$

*Resolviendo con respecto a $x$, se tiene*

$$3x = \frac{y}{p} - 6py^2$$

*Luego, derivando con respecto a $y$,*

$$\left(1 + 6p^2y\right)\left(2p + y\frac{dp}{dy}\right) = 0$$

*Igualando a cero el segundo factor se deduce que $py^2 = C$. Despejando $p$ y sustituyendo en la ecuación diferencial original se obtiene la primitiva $y^3 = 3Cx + 6C^2$. Nótese que al igualar a cero el primer factor se tiene solamente la ecuación algebraica, $1+6p^2y = 0$, con respecto a $p$. Despejando $p$ y sustituyendo en la ecuación diferencial original se obtiene la solución singular que no contiene ninguna constante arbitraria, esto es, $24y^3 + 9x^2 = 0$.*

### 1.4.4. Ecuación de Clairaut

La ecuación de la forma

$$y = px + f(p) \tag{1.54}$$

se conoce como **ecuación de Clairaut**. Su primitiva es

$$y = Cx + f(C) \tag{1.55}$$

que se comprueba de la siguiente manera. Derivando la ecuación (1.54), tenemos $p = p + x\frac{dp}{dx} + \frac{df}{dp}\frac{dp}{dx}$, de donde $\frac{dp}{dx}\left(x + \frac{df}{dp}\right) = 0$. Una solución es $\frac{dp}{dx} = 0$, de donde $p = C$. Obténgase la primitiva $y = Cx + f(C)$ sustituyendo $p = C$ en la ecuación (1.54). La ecuación $\frac{df(p)}{dp} = -x$ es algebraica y no contiene derivada de $p$ con respecto a $x$. Despejando $p$ como función de $x$ y sustituyendo en la ecuación diferencial original se obtiene una solución singular que no contiene ninguna constante arbitraria.

**Ejemplo 1.17.** *La ecuación* $y = px + \sqrt{4+p^2}$ *tiene la primitiva* $y = Cx + \sqrt{4+C^2}$. *Despejando* $p$ *de la ecuación* $-x = d\left(\sqrt{4+p^2}\right)/dp = p/\sqrt{4+p^2}$ *se tiene* $p = \pm 2x/\sqrt{1-x^2}$, *y sustituyendo en la ecuación diferencial original se obtiene la solución singular,* $y = 2\left(1 \pm x^2\right)/\sqrt{1-x^2}$.

**Ejemplo 1.18.** *La ecuación* $y = px + p^2$ *tiene la primitiva* $y = Cx + C^2$. *La solución singular es* $y = -x^2/4$.

## 1.5. EDO de orden $n$

### 1.5.1. EDO (n,1) lineales

Se dice que una ecuación diferencial de orden $n$ es lineal si tiene la forma

$$P_0(x)y^{(n)} + P_1(x)y^{(n-1)} + \cdots + P_{n-1}(x)y^{(1)} + P_n(x)y = Q(x) \tag{1.56}$$

donde $P_0(x) \neq 0, P_1(x), \ldots, P_n(x)$, $Q(x)$ son funciones de $x$ o constantes y $y^{(n)} = \frac{d^n y}{dx^n}$. Introduciendo el operador diferencial del orden $n$,

$$L^{(n)} = P_0(x)\frac{d^n}{dx^n} + P_1(x)\frac{d^{n-1}}{dx^{n-1}} + \cdots + P_n(x) \tag{1.57}$$

podemos escribir la ecuación (1.56) en forma compacta,

$$L^{(n)}\left[y\right] = Q\left(x\right) \tag{1.58}$$

La característica distintiva de esta ecuación es su linealidad con respecto a la función desconocida $y\left(x\right)$ y sus derivadas, es decir, la ecuación es de primer grado con respecto a éstas.

Si $Q\left(x\right) \equiv 0$, queda la ecuación

$$L^{(n)}\left[y\right] = 0 \tag{1.59}$$

que se llama **homogénea**; en el caso de $Q\left(x\right) \neq 0$, la ecuación (1.56) es **no homogénea**.

**Ejemplo 1.19.** *La ecuación diferencial*

$$x^2\frac{d^3y}{dx^3} - 3x\frac{d^2y}{dx^2} + 5\frac{dy}{dx} + xy = \cos x$$

*es lineal de tercer orden no homogénea, mientras que la siguiente ecuación lineal es de segundo orden homogénea,*

$$3x\frac{d^2y}{dx^2} - 5\frac{dy}{dx} + x^2y = 0$$

Las ecuaciones diferenciales lineales homogéneas poseen una importante propiedad, llamada el **principio de superposición**, que se expresa por medio del siguiente teorema.

**Teorema 1.2.** *Si $y_1(x)$, $y_2(x)$, ..., $y_m(x)$ son soluciones de una ecuación diferencial lineal homogénea, ecuación (1.59), la combinación lineal*

$$y = C_1 y_1 + C_2 y_2 + \cdots + C_m y_m$$

*donde $C_1, C_2, \ldots, C_m$ son constantes, también es la solución de la ecuación (1.59).*

**Demostración.** *Sean $y_1, y_2, \ldots, y_m$ soluciones de la ecuación (1.59), es decir,*

$$L^{(n)}[y_1] = 0, \quad L^{(n)}[y_2] = 0, \quad \ldots, \quad L^{(n)}[y_m] = 0$$

*Es fácil ver que de cualquier solución $y(x)$ de la ecuación (1.59), multiplicada por cualquier constante $C$, resulta la solución de la misma ecuación:*

$$L^{(n)}[Cy] = CL^{(n)}[y] = 0$$

*Sea $y = C_1 y_1 + C_2 y_2 + \cdots + C_m y_m$. Entonces,*

$$\begin{aligned} L^{(n)}[y] &= L^{(n)}[C_1 y_1 + C_2 y_2 + \cdots + C_m y_m] \\ &= L^{(n)}[C_1 y_1] + L^{(n)}[C_2 y_2] + \cdots + L^{(n)}[C_m y_m] \\ &= C_1 L^{(n)}[y_1] + C_2 L^{(n)}[y_2] + \cdots + C_m L^{(n)}[y_m] = 0 \end{aligned}$$

*y queda demostrado el teorema.* ■

Se dice que las funciones $y_1$, $y_2$, ... $y_m$ son **linealmente independientes** sobre un intervalo $x_1 < x < x_2$ si para toda $x$ de dicho intervalo, la igualdad

$$C_1 y_1(x) + C_2 y_2(x) + \cdots + C_m y_m(x) = 0 \tag{1.60}$$

se cumple sólo para las constantes $C_1 = C_2 = \cdots = C_m = 0$. En caso contrario las funciones se llaman **linealmente dependientes**.

**Ejemplo 1.20.** *Las funciones $y_1 = e^x$, $y_2 = e^{-x}$ son linealmente independientes. Para demostrarlo supóngase la relación*

$$C_1 e^x + C_2 e^{-x} = 0 \Longleftrightarrow C_1 y_1 + C_2 y_2 = 0$$

*donde $C_1$ y $C_2$ son constantes. Al derivarla se obtiene*

$$C_1 e^x - C_2 e^{-x} = 0 \Longleftrightarrow C_1 y_1' + C_2 y_2' = 0$$

*Las dos ecuaciones forman el sistema homogéneo de ecuaciones lineales con respecto a $C_1$ y $C_2$. Resolviéndolo se tiene $C_1 = C_2 = 0$.*

**Ejemplo 1.21.** *Las funciones* $y_1 = e^x$, $y_2 = e^{-x}$, $y_3 = 7e^x$ *son linealmente dependientes, ya que para* $C_1 = 7, C_2 = 0, C_3 = -1$ *se cumple la igualdad* $C_1e^x + C_2e^{-x} + 7C_3e^x = 0$ *para toda* $x$.

El concepto de independencia lineal de funciones tiene su importancia al aplicarlo a las soluciones de ecuaciones diferenciales lineales. El criterio de independencia lineal se establece en términos del **determinante de Wronski** o, simplemente, **wronskiano**:

$$W(y_1, y_2, \cdots, y_m) = \begin{vmatrix} y_1 & y_2 & \cdots & y_m \\ y_1' & y_2' & \cdots & y_m' \\ \vdots & \vdots & & \vdots \\ y_1^{(m-1)} & y_2^{(m-1)} & \cdots & y_m^{(m-1)} \end{vmatrix} \tag{1.61}$$

donde las funciones $y_1(x), y_2(x), \ldots, y_m(x)$ tienen derivadas continuas con respecto a $x$ hasta el orden $(m-1)$.

**Teorema 1.3.** *(La condición necesaria y suficiente) Las funciones* $y_1(x)$, $y_2(x)$, $\ldots$, $y_m(x)$, *que tienen sus derivadas continuas hasta el orden* $(m-1)$, *son linealmente independientes sobre un intervalo* $x_1 < x < x_2$ *si y sólo si su wronskiano* $W(y_1, y_2, \ldots, y_m) \neq 0$ *en este intervalo. (Si* $W = 0$ *para un valor* $x = x_0$, *entonces* $W \equiv 0$ *sobre todo el intervalo).*

**Demostración.** *Supongamos que las funciones* $y_1(x), y_2(x), \ldots, y_m(x)$ *son linealmente dependientes, pues existe una combinación lineal tal que* $C_1y_1(x) + C_2y_2(x) + \cdots + C_my_m(x) = 0$. *Derivando a ésta* $(m-1)$ *veces consecutivamente se obtiene el sistema homogéneo de ecuaciones lineales con respecto a las constantes incógnitas* $C_1, C_2, \ldots, C_m$ :

$$\begin{gathered} C_1y_1(x) + C_2y_2(x) + \cdots + C_my_m(x) = 0 \\ C_1y_1'(x) + C_2y_2'(x) + \cdots + C_my_m'(x) = 0 \\ \vdots \\ C_1y_1^{(m-1)}(x) + C_2y_2^{(m-1)}(x) + \cdots + C_my_m^{(m-1)}(x) = 0 \end{gathered}$$

*Dicho sistema tiene una solución no trivial si y sólo si su determinante, que es precisamente el wronskiano, es igual a cero. Lo último contradice la condición del teorema de que* $W(y_1, y_2, \ldots, y_m) \neq 0$. *Por tanto, la suposición de la dependencia lineal de las funciones* $y_1, y_2, \ldots, y_m$ *es errónea. Queda probado el teorema.* ■

**Proposición 1.6.** *La ecuación diferencial lineal homogénea de orden* $n$, *ecuación (1.59), tiene no más que* $n$ *soluciones linealmente independientes. Si* $y_1(x)$, $y_2(x)$, $\ldots$, $y_n(x)$ *son* $n$ *soluciones linealmente independientes de la ecuación (1.59), entonces, la primitiva de la ecuación (1.59) es*

$$y_0(x) = C_1y_1(x) + C_2y_2(x) + \cdots + C_ny_n(x) \tag{1.62}$$

*donde* $C_1, C_2, \ldots, C_n$ *son constantes arbitrarias.*

Un conjunto de $n$ soluciones linealmente independientes $y_1(x), y_2(x), \ldots, y_n(x)$, se llama **base** o bien, **sistema fundamental de soluciones** de la ecuación (1.59).

**Proposición 1.7.** *Sea $y_p(x)$ cualquier solución particular de la ecuación diferencial lineal no homogénea, ecuación (1.56). Entonces,*

$$y(x) = C_1y_1(x) + C_2y_2(x) + \cdots + C_ny_n(x) + y_p(x) = y_0(x) + y_p(x) \tag{1.63}$$

*es la primitiva de la ecuación (1.56). Es decir, la solución general de una ecuación diferencial lineal no homogénea es la suma de una solución particular cualquiera de la ecuación no homogénea y de la solución general de la ecuación diferencial lineal homogénea.*

Aunque en general la primitiva de una ecuación diferencial no es necesariamente la solución completa de la ecuación, sin embargo, cuando la ecuación es lineal, la primitiva es su solución completa.

## 1.5.2. Ecuaciones lineales homogéneas con coeficientes constantes

Una ecuación diferencial lineal homogénea con coeficientes constantes $P_0 = 1, P_1, \ldots, P_n$ tiene la forma

$$y^{(n)} + P_1y^{(n-1)} + \cdots + P_{n-1}y^{(1)} + P_ny = 0 \tag{1.64}$$

Mediante la notación $D \equiv \frac{d}{dx}$, la ecuación (1.64) toma la forma

$$\left(D^n + P_1D^{n-1} + \cdots + P_{n-1}D + P_n\right)[y] = 0 \tag{1.65}$$

donde el operador diferencial $L^{(n)}(D) = D^n + P_1D^{n-1} + \cdots + P_{n-1}D + P_n$ tiene la forma de un polinomio en la variable $D$. Según el teorema fundamental del álgebra, un polinomio de orden $n$ tiene $n$ raíces, sean $\lambda_1, \lambda_2, \ldots, \lambda_n$, algunas de ellas pueden ser múltiples. Esto implica que el polinomio puede ser factorizado, o bien presentado en la forma

$$\begin{aligned} L^{(n)}(D) &= D^n + P_1D^{n-1} + \cdots + P_{n-1}D + P_n \\ &= (D-\lambda_1)(D-\lambda_2)\cdots(D-\lambda_n) \end{aligned}$$

Entonces la ecuación (1.65) toma la forma

$$(D-\lambda_1)(D-\lambda_2)\cdots(D-\lambda_n)[y] = 0 \tag{1.66}$$

Nótese que en el operador anterior el orden de los factores no tiene importancia, es decir, $(D-\lambda_k)(D-\lambda_m) = (D-\lambda_m)(D-\lambda_k)$.

Es fácil ver que la ecuación diferencial $(D-\lambda)[y] = 0$ tiene la solución $y = e^{\lambda x}$. Esto nos sugiere la solución de la ecuación (1.66), o bien, de la

ecuación (1.65) en forma exponencial, $y = e^{\lambda x}$. Sustituyéndola en la ecuación (1.65), se tiene

$$e^{\lambda x}\left(\lambda^n + P_1\lambda^{n-1} + \cdots + P_{n-1}\lambda + P_n\right) = 0$$

que se cumple si

$$\lambda^n + P_1\lambda^{n-1} + \cdots + P_{n-1}\lambda + P_n = 0 \qquad (1.67)$$

La última se conoce como **ecuación característica** de la ecuación (1.65); sus raíces o **valores característicos** $\lambda_1, \lambda_2, \ldots, \lambda_n$ permiten escribir la ecuación característica en la forma

$$(\lambda - \lambda_1)\,(\lambda - \lambda_2)\,\cdots\,(\lambda - \lambda_n) = 0$$

Del álgebra se sabe que la ecuación característica puede tener raíces de dos tipos:

**Caso I.** La ecuación característica tiene $n$ raíces distintas $\lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \cdots \neq \lambda_n$. Entonces las $n$ soluciones linealmente independientes

$$y_1 = e^{\lambda_1 x}, \quad y_2 = e^{\lambda_2 x}, \quad \ldots, \quad y_n = e^{\lambda_n x}$$

constituyen una base y la solución general de la ecuación (1.65), o bien de la ecuación (1.64) es

$$y_0 = C_1 e^{\lambda_1 x} + C_2 e^{\lambda_2 x} + \cdots + C_n e^{\lambda_n x} \qquad (1.68)$$

**Caso II.** Algunos de los valores característicos son múltiples. Digamos que $\lambda_1$ es una raíz múltiple de grado $k$ y las demás son distintas, $\lambda_1 \neq \lambda_{k+1} \neq \cdots \neq \lambda_n$. En este caso, las $k$ soluciones linealmente independientes que corresponden al valor característico $\lambda_1$ son

$$y_1 = e^{\lambda_1 x}, \quad y_2 = xe^{\lambda_1 x}, \quad y_3 = x^2 e^{\lambda_1 x}, \quad \ldots, \quad y_k = x^{k-1} e^{\lambda_1 x}$$

que se comprueba notando que

$$L^{(n)}(D) = (D - \lambda_1)^k\,(D - \lambda_{k+1})\cdots(D - \lambda_n)$$

y, además,

$$\begin{aligned}
(D - \lambda_1)\left[x^m e^{\lambda_1 x}\right] &= mx^{m-1}e^{\lambda_1 x}, \\
(D - \lambda_1)^2\left[x^m e^{\lambda_1 x}\right] &= m\,(m-1)\,x^{m-2}e^{\lambda_1 x} \\
&\ \ \vdots \\
(D - \lambda_1)^m\left[x^m e^{\lambda_1 x}\right] &= m!e^{\lambda_1 x} \\
(D - \lambda_1)^{m+1}\left[x^m e^{\lambda_1 x}\right] &= 0
\end{aligned}$$

De la última relación se sigue que

$$(D-\lambda_1)^k\left[x^m e^{\lambda_1 x}\right]=0 \quad \text{si} \quad m=0,1,\ldots,k-1$$

Soluciones linealmente independientes correspondientes a los otros valores característicos son

$$y_{k+1}=e^{\lambda_{k+1}x}, \quad \ldots, \quad y_n=e^{\lambda_n x}$$

y, por tanto, la solución general de la ecuación (1.65), o bien de la ecuación (1.64) es

$$y_0=C_1e^{\lambda_1 x}+C_2xe^{\lambda_1 x}+\cdots+C_kx^{k-1}e^{\lambda_1 x}+C_{k+1}e^{\lambda_{k+1}x}+\cdots+C_ne^{\lambda_n x} \quad (1.69)$$

**Ejemplo 1.22.** *En caso de una ecuación diferencial lineal de segundo orden con coeficientes constantes reales*

$$y''+2ay'+by=0$$

*la ecuación característica es*

$$\lambda^2+2a\lambda+b=0$$

*y sus raíces son*

$$\lambda_1=-a+\sqrt{a^2-b}, \quad \lambda_2=-a-\sqrt{a^2-b}$$

*a) Raíces reales distintas,* $\lambda_1 \neq \lambda_2$*: la solución general es*

$$y_0=C_1e^{\lambda_1 x}+C_2e^{\lambda_2 x}$$

*b) Raíces complejas conjugadas,* $\lambda_1=-a+i\beta$, $\lambda_2=-a-i\beta$*, donde* $\beta=\sqrt{b-a^2}$*. Para este caso,*

$$y_1=e^{\lambda_1 x}=e^{-\alpha x}e^{i\beta x}, \qquad y_2=e^{\lambda_2 x}=e^{-\alpha x}e^{-i\beta x}$$

*y tomando* $C_2=C_1^*$ *la solución general real es*

$$y_0=C_1e^{\lambda_1 x}+C_1^*e^{\lambda_2 x}=Ae^{-\alpha x}\cos(\beta x+\phi)$$

*Se queda como tarea el cálculo de las constantes arbitrarias* $A$ *y* $\phi$ *en función de* $C_1$.

*c) Una raíz real doble,* $\lambda=-a$, $(a^2=b)$*: la solución general es*

$$y_0=C_1e^{\lambda x}+C_2xe^{\lambda x}=C_1e^{-ax}+C_2xe^{-ax}$$

### 1.5.3. EDO (n,1) lineales no homogéneas

La solución general $y(x)$ de una ecuación diferencial lineal

$$L^{(n)}[y] = Q(x) \tag{1.70}$$

es la suma $y(x) = y_0 + y_p$ de una solución particular $y_p$ de esta ecuación y de la solución general $y_0$ de la ecuación homogénea correspondiente. Se conocen dos métodos para obtener $y_p$. El **método de variación de parámetros** es completamente general. El **método de los coeficientes indeterminados**, siendo un método especial y más sencillo, sin embargo es práctico para aplicaciones en problemas de ingeniería.

**Coeficientes indeterminados**

Este método es aplicable a ecuaciones lineales con coeficientes constantes cuando $Q(x)$ tiene la forma

$$Q(x) = \left(Ae^{\alpha x} \pm Be^{\beta x}\right) P_n(x)$$

en donde $A$ y $B$ son algunas constantes y

$$P_n(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_n$$

Para obtener una solución particular de la ecuación (1.70) se supone $y_p$ de la forma semejante a la de $Q(x)$ y sus derivadas, conteniendo coeficientes desconocidos que se determinan sustituyendo $y_p$ y sus derivadas en la ecuación (1.70). A continuación se presentan ejemplos ilustrativos.

**Ejemplo 1.23.** $y''' + ay = 3x$. *Suponiendo que* $y_p = Ax + B = A_1Q(x) + B_1Q'(x)$ *y sustituyendo en la ecuación diferencial, se tiene* $a(Ax+B) = 3x$, *de donde* $a(Ax) = 3x$, *pues* $A = \frac{3}{a}$; *también,* $aB = 0$, *y* $B = 0$. *Así,* $y_p = \frac{3}{a}x$.

**Ejemplo 1.24.** $y''' + ay = e^{bx}$. *Intentaremos* $y_p = Ae^{bx}$, *por lo que* $Ab^3e^{bx} + aAe^{bx} = e^{bx}$, *luego* $Ab^3 + Aa = 1$, *de donde* $A = \left(a+b^3\right)^{-1}$, *obteniéndose* $y_p = \left(a+b^3\right)^{-1} e^{bx}$.

**Ejemplo 1.25.** $y'' - 2y' + 3y = x^3 + \operatorname{sen} x$. *Suponemos que* $y_p = A_3x^3 + A_2x^2 + A_1x + A_0 + B\operatorname{sen} x + C\cos x$. *Sustituyendo, se obtiene*

$$\begin{aligned} x^3 + \operatorname{sen} x &= 6A_3x + 2A_2 - B\operatorname{sen} x - C\cos x - \\ &\quad 2\left(3A_3x^2 + 2A_2x + A_1 + B\cos x - C\operatorname{sen} x\right) + \\ &\quad 3\left(A_3x^3 + A_2x^2 + A_1x + A_0 + B\operatorname{sen} x + C\cos x\right) \end{aligned}$$

*Igualando los coeficientes de* $\cos x$ *y* $\operatorname{sen} x$, *y de potencias de* $x$, *se tiene*

$$\begin{aligned} -C - 2B + 3C &= 0, \quad -B + 2C + 3B = 1 \\ 3A_3 &= 1, \quad -6A_3 + 3A_2 = 0 \\ 6A_3 - 4A_2 + 3A_1 &= 0, \quad 2A_2 - 2A_1 + 3A_0 = 0 \end{aligned}$$

*De donde,* $A_3 = 1/3$, $A_2 = 2/3$, $A_1 = 2/9$, $A_0 = -8/27$, $B = C = 1/4$ *y, por lo tanto,*

$$y_p = \frac{1}{3}x^3 + \frac{2}{3}x^2 + \frac{2}{9}x - \frac{8}{27} + \frac{1}{4}\operatorname{sen} x + \frac{1}{4}\cos x$$

**Solución por variación de parámetros**

Para entender mejor la idea del método empezaremos con ecuaciones diferenciales lineales de segundo orden de la forma

$$y'' + A(x)\,y' + B(x)\,y = Q(x) \tag{1.71}$$

La ecuación diferencial homogénea correspondiente

$$y'' + A(x)\,y' + B(x)\,y = 0 \tag{1.72}$$

tiene una solución general $y_0(x)$ de la forma

$$y_0(x) = C_1 y_1(x) + C_2 y_2(x)$$

donde $y_1(x)$ y $y_2(x)$ son dos soluciones linealmente independientes de la ecuación (1.72), y $C_1$ y $C_2$ son dos constantes arbitrarias.

El método de variación de parámetros consiste en reemplazar las constantes $C_1$ y $C_2$ por las funciones $C_1(x)$ y $C_2(x)$ que deben determinarse de modo que la función resultante

$$y_p(x) = C_1(x)\,y_1(x) + C_2(x)\,y_2(x) \tag{1.73}$$

sea una solución particular de la ecuación (1.71). Derivando se tiene

$$y_p' = C_1 y_1' + C_2 y_2' + C_1' y_1 + C_2' y_2$$

La $y_p(x)$ contiene dos funciones $C_1(x)$ y $C_2(x)$, pero la condición de que $y_p$ satisfaga la ecuación (1.71) impone sólo una relación entre ellas. Teniendo la libertad de imponer una segunda condición, suponemos que

$$C_1' y_1 + C_2' y_2 = 0 \tag{1.74}$$

Entonces,

$$y_p' = C_1 y_1' + C_2 y_2' \tag{1.75}$$

$$y_p'' = C_1 y_1'' + C_2 y_2'' + C_1' y_1' + C_2' y_2' \tag{1.76}$$

Sustituyendo (1.75) y (1.76) en la ecuación (1.71), se obtiene

$$\begin{aligned} &(C_1 y_1'' + C_2 y_2'') + C_1' y_1' + C_2' y_2' + A(x)\,(C_1 y_1' + C_2 y_2') + \\ &B(x)\,(C_1 y_1 + C_2 y_2) \\ = \ &Q(x) \end{aligned}$$

Como $y_1(x)$ y $y_2(x)$ son soluciones de la ecuación homogénea, ecuación (1.72), esta relación junto con la ecuación (1.74) se reduce al sistema algebraico de dos ecuaciones lineales con respecto a dos funciones incógnitas $C_1'$ y $C_2'$,

$$\begin{aligned} C_1'y_1 + C_2'y_2 &= 0 \\ C_1'y_1' + C_2'y_2' &= Q(x) \end{aligned}$$

en donde el determinante del sistema es el wronskiano de $y_1$ y $y_2$, $W(y_1, y_2) = y_1y_2' - y_1'y_2 \neq 0$. Por la regla de Cramer se halla

$$C_1' = -\frac{y_2Q}{W}, \quad C_2' = \frac{y_1Q}{W}$$

Integrando

$$C_1(x) = -\int \frac{y_2Q}{W}dx, \quad C_2(x) = \int \frac{y_1Q}{W}dx$$

y sustituyendo estas expresiones en la ecuación (1.73), se obtiene la solución particular deseada de la ecuación (1.71),

$$y_p(x) = -y_1(x)\int \frac{y_2Q}{W}dx + y_2(x)\int \frac{y_1Q}{W}dx \tag{1.77}$$

Si se dejan las constantes arbitrarias de integración que aparecen en la ecuación (1.77), entonces ésta se convierte en una solución general de la ecuación (1.71).

La generalización del método de variación de parámetros para una ecuación diferencial lineal no homogénea de orden $n$, ecuación (1.70), es bastante directa. La solución particular se busca en la forma

$$y_p(x) = C_1(x)y_1(x) + C_2(x)y_2(x) + \cdots + C_n(x)y_n(x)$$

donde $y_1(x), y_2, \ldots, y_n$ son soluciones fundamentales de la ecuación homogénea correspondiente. La $y_p(x)$ contiene $n$ funciones $C_1(x)$, $C_2(x)$, $\ldots$, $C_n(x)$ desconocidas, pero la condición de que $y_p$ satisfaga la ecuación (1.70) impone sólo una relación entre ellas. Teniendo la libertad de imponer $(n-1)$ condiciones adicionales, suponemos que

$$\begin{gathered} C_1'y_1 + C_2'y_2 + \cdots + C_n'y_n = 0 \\ C_1'y_1' + C_2'y_2' + \cdots + C_n'y_n' = 0 \\ C_1'y_1'' + C_2'y_2'' + \cdots + C_n'y_n'' = 0 \\ \vdots \\ C_1'y_1^{(n-2)} + C_2'y_2^{(n-2)} + \cdots + C_n'y_n^{(n-2)} = 0 \end{gathered}$$

Junto con la relación

$$C_1'y_1^{(n-1)} + C_2'y_2^{(n-1)} + \cdots + C_n'y_n^{(n-1)} = Q(x)$$

que se obtiene al sustituir $y_p$ en la ecuación (1.70), éstas constituyen un sistema algebraico de $n$ ecuaciones lineales con respecto a $C_1'$, $C_2'$, ..., $C_n'$, que es resuelto por la regla de Cramer, dado que su determinante es wronskiano $W(y_1, y_2, \ldots, y_n) \neq 0$. Luego, integrando se obtienen $C_1(x)$, $C_2(x)$, ..., $C_n(x)$ y finalmente $y_p(x)$.

**Ejercicio.** *Obtener las primitivas de las ecuaciones*

$$\begin{aligned}(D^2 - 6D + 9)y &= x^{-2}e^{3x} \\ (D^2 - 2D + 3)y &= x^3 + \operatorname{sen} x\end{aligned}$$

*por variación de parámetros.*

## 1.6. Problemas con valores frontera e iniciales

Cada ecuación diferencial ordinaria tiene sólo una variable independiente, la cual, a menudo, es el tiempo. En muchas aplicaciones se tiene interés en una solución particular $y(t)$ de una ecuación diferencial dada. La solución general $y(t; C_1, C_2, \ldots, C_n)$ de una ecuación diferencial ordinaria de orden $n$ contiene un número $n$ de constantes arbitrarias, $C_1$, $C_2$, ..., $C_n$. Por tanto, para definir una solución particular de esta ecuación hay que dar ciertos valores a estas constantes. En la práctica los valores de $n$ constantes arbitrarias se determinan de $n$ condiciones que se conocen como **condiciones iniciales**, dado que éstas prescriben la situación inicial a un problema físico en un instante, digamos $t = 0$. En muchas aplicaciones son $n$ condiciones del tipo

$$\begin{aligned}y(t; C_1, C_2, \ldots, C_n)|_{t=0} &= K_0, \quad y'(t; C_1, C_2, \ldots, C_n)|_{t=0} = K_1, \ldots, \\ y^{(n-1)}(t; C_1, C_2, \ldots C_n)|_{t=0} &= K_{n-1}\end{aligned}$$

donde $K_0, K_2, \ldots, K_{n-1}$ son números dados. Una ecuación diferencial ordinaria junto con sus condiciones iniciales se llama **problema con valor inicial**.

Las ecuaciones diferenciales parciales (en derivadas parciales) tienen por lo menos dos variables independientes. Por razones físicas, un conjunto de variables independientes suele separarse en una variable temporal, el tiempo $t$, y en un grupo de variables espaciales, por ejemplo $x, y, z$. Dicha diferencia entre las ecuaciones ordinarias y parciales da origen a dos tipos de **condiciones, iniciales** y **en la frontera**, para definir una solución particular de la ecuación diferencial parcial.

Una ecuación muy conocida es la ecuación de onda que describe la propagación de una perturbación (una onda) en un medio unidimensional, por ejemplo en una cuerda elástica:

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad t > 0, \quad 0 < x < L$$

En esta ecuación se supone que la cuerda de longitud $L$ está ubicada sobre el eje $x$ de un sistema de coordenadas con un extremo en el punto $x = 0$ y con

el otro en $x = L$. La solución $u(x,t)$ describe la propagación de onda a partir del instante $t = 0$. Para definir una solución particular se requieren condiciones iniciales y condiciones en la frontera. Siendo una ecuación de segundo orden con respecto a la derivada temporal, hay que aplicar dos condiciones iniciales que son los valores de la solución y de su primera derivada temporal en todo el intervalo espacial:

$$u(x,0) = a_1(x), \; \frac{\partial u}{\partial t}|_{t=0} = a_2(x), \quad 0 < x < L$$

donde $a_1(x)$ y $a_2(x)$ son funciones dadas en el intervalo $0 < x < L$. Dado que la ecuación de onda es de segundo orden con respecto a la derivada espacial, hay que satisfacer dos condiciones en la frontera, por ejemplo,

$$u(0,t) = b_1(t), \quad u(L,t) = b_2(t), \quad t > 0$$

que son valores de la onda en los extremos de la cuerda, donde $b_1(t)$ y $b_2(t)$ son funciones dadas de tiempo. Siendo siempre los dos para la ecuación de onda, condiciones de la frontera dependiendo de las condiciones físicas que los determinan, pueden tener una forma diferente a los escritos anteriormente. Por ejemplo, podrían ser una combinación lineal de la función y su primera derivada espacial en los dos extremos de la cuerda: $\alpha_1 u(0,t) + \beta_1 \frac{\partial u(x,t)}{\partial x}|_{x=0} = b_1(t), \quad \alpha_2 u(L,t) + \beta_2 \frac{\partial u(x,t)}{\partial x}|_{x=L} = b_2(t), \quad t > 0$. Una ecuación diferencial parcial junto con sus condiciones iniciales y en la frontera se llama **problema con condiciones frontera**.

## 1.7. Reducción de un sistema general a un sistema normal

Los teoremas de existencia y unicidad de soluciones se formulan, por lo general, en términos de sistemas normales de ecuaciones diferenciales. Un sistema general de EDO es muy común tanto en modelos matemáticos de ciencias naturales y exactas como en métodos numéricos. Dada la importancia de los sistemas de EDO, en esta sección se considera la reducción de una EDO o un sistema general de EDO a un sistema normal.

Un sistema de EDO en forma

$$u_i' = f_i(x, u_1, u_2, \ldots, u_n), \quad i = 1, 2, \ldots, n \tag{1.78}$$

se llama **sistema normal**. En este sistema $x$ es la variable independiente, $u_1, u_2, \ldots, u_n$ son funciones incógnitas de la variable $x$, $u_i' = du_i/dx$, $f_1, f_2, \ldots, f_n$ son funciones dadas de las variables $x, u_1, u_2, \ldots, u_n$ en un dominio abierto del espacio de $(n+1)$ dimensiones. Se supone que las funciones $f_1, f_2, \ldots, f_n$ son continuas junto con sus derivadas parciales,

$$\frac{\partial f_i(x, u_1, u_2, \ldots, u_n)}{\partial u_j}, \quad i, j = 1, 2, \ldots, n$$

en el dominio de su definición. Un caso particular de sistemas normales es un sistema normal de EDO lineales:

$$u_i' = \sum_{j=1}^{n} a_{ij}(x)\, u_j + b_i(x), \quad i = 1, 2, \ldots, n \tag{1.79}$$

o en forma vectorial,

$$\frac{d\mathbf{u}}{dx} = A(x)\,\mathbf{u} + \mathbf{b}(x) \tag{1.80}$$

en donde la matriz $A(x) = [a_{ij}(x)]$ y el vector $\mathbf{b}(x)$ son funciones dadas de la variable independiente $x$.

En la práctica surgen ecuaciones y sistemas de EDO más generales que lo escrito en párrafos anteriores. En el caso de una sola función incógnita $y(x)$ de una variable independiente $x$, la ecuación diferencial de orden $n$ puede escribirse en la forma

$$F\left(x, y, y', \ldots, y^{(n)}\right) = 0 \tag{1.81}$$

en donde $F$ es una función dada de $(n+2)$ variables que está definida en un dominio abierto del espacio de sus variables. Una función $y = \varphi(x)$ definida en un intervalo $x_1 < x < x_2$, que tiene las derivadas hasta el orden $n$, es la solución de la ecuación (1.81) si al sustituirla en ésta se obtiene la identidad en la variable $x$ en el intervalo $x_1 < x < x_2$.

En el caso de dos funciones incógnitas, $y(x)$ y $z(x)$, de una sola variable independiente $x$, que no son independientes, se consideran dos ecuaciones diferenciales que forman un sistema. Este sistema se puede escribir en la forma

$$\begin{aligned} F\left(x, y, y', \ldots, y^{(n)}, z, z', \ldots, z^{(m)}\right) = 0 \\ G\left(x, y, y', \ldots, y^{(n)}, z, z', \ldots, z^{(m)}\right) = 0 \end{aligned} \tag{1.82}$$

en donde $F$ y $G$ son funciones dadas, cada una $(n+m+3)$ variables, definidas en un dominio abierto. Si el orden de la derivada mayor de la función $y(x)$ es $n$ y el orden de la derivada mayor de $z(x)$ es $m$, se dice que el sistema (1.82) es de orden $n$ con respecto a $y(x)$ y de orden $m$ con respecto a $z(x)$. El número $(n+m)$ es el orden del sistema (1.82). La solución de (1.82) es un par de funciones $y = \varphi(x)$ y $z = \psi(x)$, definidas en un intervalo $x_1 < x < x_2$, tales que al sustituirlas en (1.82) se obtienen dos identidades en $x$ en el intervalo $x_1 < x < x_2$.

De modo similar se define un sistema de EDO de tres y más funciones incógnitas de una sola variable. Sean $y_1, y_2, \ldots, y_n$ funciones incógnitas y sea $m_i$ el orden de derivada mayor de $y_i$, $i = 1, 2, \ldots, n$. Entonces, se dice que $m_i$ es el orden del sistema con respecto a $y_i$ y $m = m_1 + m_2 + \cdots + m_n$ es el orden del sistema. Por tanto, el sistema normal (1.78) tiene el orden $n$.

Si la ecuación (1.81) puede ser despejada con respecto a la derivada mayor, entonces esta ecuación se escribe en la forma

$$y^{(n)} = f\left(x, y, y', \ldots, y^{(n-1)}\right) \tag{1.83}$$

De manera similar, si las ecuaciones del sistema (1.82) pueden ser resueltas con respecto a las derivadas mayores, entonces este sistema se escribe en la forma

$$\begin{aligned} y^{(n)} &= f\left(x, y, y', \ldots, y^{(n-1)}, z, z', \ldots, z^{(m-1)}\right) \\ z^{(m)} &= g\left(x, y, y', \ldots, y^{(n-1)}, z, z', \ldots, z^{(m-1)}\right) \end{aligned} \tag{1.84}$$

Las ecuaciones (1.83) y (1.84) se llaman **resueltas con respecto a las derivadas mayores**. De manera similar se definen sistemas de ecuaciones con un número arbitrario de funciones incógnitas que son resueltas con respecto a las derivadas mayores. En particular, el sistema normal, ecuación (1.78), es resuelto con respecto a las derivadas mayores.

En lo que sigue se demuestra que cualquier sistema de EDO de orden $n$ resuelto con respecto a las derivadas mayores es reducible a un sistema normal de orden $n$. Para empezar, mostramos que una ecuación diferencial de orden $n$ se reduce a un sistema normal de orden $n$. La ecuación (1.83) es una ecuación diferencial de orden $n$ resuelta con respecto a la derivada mayor, en donde $f\left(x, y, y', \ldots, y^{(n-1)}\right)$ es una función dada de $(n+1)$ variables en un dominio abierto. Se supone que tanto $f\left(x, y, y', \ldots, y^{(n-1)}\right)$ como sus derivadas parciales $\partial f/\partial y^{(k)}$, $k = 0, 1, \ldots, n-1$, son continuas en el dominio de definición de $f$. Con el fin de transformar la ecuación (1.83) en un sistema normal se introducen nuevas funciones incógnitas, $u_1, u_2, \ldots, u_n$, de la variable independiente $x$:

$$u_1 = y, \quad u_2 = y', \quad \ldots, \quad u_n = y^{(n-1)} \tag{1.85}$$

Se tiene que la ecuación (1.83) es equivalente al siguiente sistema normal:

$$\begin{aligned} u_1' &= u_2 \\ u_2' &= u_3 \\ &\vdots \\ u_{n-1}' &= u_n \\ u_n' &= f\left(x, u_1, u_2, \ldots, u_n\right) \end{aligned} \tag{1.86}$$

El método descrito en el párrafo anterior permite reducir a la forma normal cualquier sistema de EDO resuelto con respecto a las derivadas mayores. Con el fin de no complicar las expresiones, consideremos el sistema de cuarto orden constituido por dos ecuaciones de segundo orden:

$$\begin{aligned} y'' &= f\left(x, y, y', z, z'\right) \\ z'' &= g\left(x, y, y', z, z'\right) \end{aligned} \tag{1.87}$$

Dicho sistema se reduce a la forma normal por medio del cambio a las nuevas funciones incógnitas:

$$u_1 = y, \quad u_2 = y', \quad u_3 = z, \quad u_4 = z' \tag{1.88}$$

Finalmente, se tiene que el sistema (1.87) es equivalente al siguiente sistema normal:

$$\begin{aligned} u_1' &= u_2 \\ u_2' &= f\left(x, u_1, u_2, u_3, u_4\right) \\ u_3' &= u_4 \\ u_4' &= g\left(x, u_1, u_2, u_3, u_4\right) \end{aligned} \tag{1.89}$$

**Nota bibliográfica:** para ampliar los conocimientos sobre el tema de este capítulo, se puede consultar [2], [4], [5] y [11].

# Capítulo 2

# Matrices

## 2.1. Definición y origen de las matrices

Tanto en las matemáticas puras como en las aplicadas surgen múltiples situaciones en las que tenemos que tratar con arreglos rectangulares de números o funciones. Algunos de estos ejemplos son los siguientes.

**Ejemplo 2.1.** *Un número complejo $x+iy$, donde $x$ y $y$ son números reales e $i$ es una unidad imaginaria, se define completamente por el par ordenado $(x,y)$ de números reales, es decir, por el arreglo de $1\times 2$ elementos.*

**Ejemplo 2.2.** *Las coordenadas cartesianas de un punto en el espacio tridimensional constituyen un triplete ordenado $(x,y,z)$ de números reales $x$, $y$ y $z$.*

**Ejemplo 2.3.** *La forma cuadrática*

$$ax^2+by^2+cz^2+2hxy+2gxz+2fyz$$

*está completamente determinada por el arreglo simétrico de $3\times 3$ de números*

$$\begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix}$$

**Ejemplo 2.4.** *El sistema de ecuaciones lineales*

$$\begin{aligned} a_{11}x_1+a_{12}x_2+a_{13}x_3&=c_1\\ a_{21}x_1+a_{22}x_2+a_{23}x_3&=c_2\\ a_{31}x_1+a_{32}x_2+a_{33}x_3&=c_3 \end{aligned}$$

*está completamente determinado por el arreglo de $3\times 4$ de coeficientes*

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & c_1 \\ a_{21} & a_{22} & a_{23} & c_2 \\ a_{31} & a_{32} & a_{33} & c_3 \end{bmatrix}$$

*dado que toda la información sobre las soluciones de este sistema de ecuaciones lineales se puede expresar en términos de propiedades del arreglo.*

**Ejemplo 2.5.** *La transformación lineal*

$$u = a_{11}x_1 + a_{12}x_2 + a_{13}x_3$$
$$v = a_{21}x_1 + a_{22}x_2 + a_{23}x_3$$

*que mapea los puntos* $(x_1, x_2, x_3)$ *del espacio tridimensional en los puntos* $(u, v)$ *de un plano (espacio bidimensional), se determina completamente por el arreglo de* $2 \times 3$ *de números*

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

*Dicha transformación lineal se puede estudiar en términos del álgebra de arreglos rectangulares correspondientes.*

La generalización de los ejemplos presentados anteriormente se da por medio del concepto de matriz.

**Definición 2.1.** *Una disposición (un arreglo) rectangular*

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \tag{2.1}$$

*de* $m \times n$ *elementos tomados de un conjunto de elementos* $S$ *y ordenados en* $m$ *renglones y* $n$ *columnas se llama* **matriz** *de orden* $m \times n$ *sobre el conjunto* $S$*, si uno quiere subrayar la naturaleza de los elementos de la matriz* $A$*. Los elementos del conjunto* $S$ *se llaman escalares. Si el conjunto* $S$ *representa todos los números complejos o reales, las matrices definidas sobre estos conjuntos se llaman matrices complejas o reales, respectivamente.*

Cuando $m = n$, la matriz se llama **cuadrada** de orden $n$ y se dice que los elementos $a_{11}, a_{22}, \ldots, a_{nn}$ constituyen la diagonal principal de $A$, mientras que los elementos $a_{n1}, a_{n-1,2}, \ldots, a_{1n}$ constituyen la diagonal secundaria. Se emplean paréntesis rectangulares o barras dobles para representar matrices, es decir, $A = [a_{ij}] = ||a_{ij}||$, donde los subíndices $i$ y $j$ del elemento $a_{ij}$ de la matriz $A$ identifican, respectivamente, el renglón y la columna en los que está situado $a_{ij}$. Casos particulares de matrices son el **vector columna** $A$, es decir, la matriz de orden $m \times 1$

$$A = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} \equiv \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}$$

y el **vector renglón** $B$, que es la matriz de orden $1 \times n$

$$B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \end{bmatrix}$$

**Ejercicio.** *Dada la matriz*

$$A = \begin{bmatrix} 1 & 5i & 7 \\ 3i & -1 & -6 \end{bmatrix}$$

*a) ¿Cuáles son las dimensiones de $A$?; b) ¿a qué equivalen los elementos $a_{12}$, $a_{21}$, $a_{22}$?*

**Ejercicio.** *Escribir en forma explícita la matriz $A = [a_{ij}]$ determinada por las siguientes condiciones: a) $m = 3$, $n = 2$, $a_{ij} = |2i - j|$; b) $m = n = 3$, $a_{ij} = |i - j|$.*

## 2.2. Operaciones básicas

En esta sección introducimos operaciones fundamentales sobre matrices, que son la adición de matrices, la multiplicación por un escalar, la multiplicación de matrices y la matriz inversa (de la izquierda, de la derecha) y la derivada de una matriz de funciones. Ante todo empezamos por la igualdad de matrices.

### 2.2.1. Igualdad de matrices

**Definición 2.2.** *Se dice que dos matrices $A = [a_{ij}]$ y $B = [b_{ij}]$ son iguales, $A = B$, si y sólo si tienen el mismo orden y todos sus elementos correspondientes son iguales, $a_{ij} = b_{ij}$.*

De esta definición se siguen cuatro propiedades de la **igualdad de matrices**:

1) Si $A$ y $B$ son dos matrices cualesquiera, se tiene que $A = B$ o $A \neq B$ (propiedad determinativa).

2) Si $A$ es una matriz cualquiera, entonces $A = A$ (propiedad reflexiva).

3) Si $A = B$, entonces $B = A$ (propiedad simétrica).

4) Si $A = B$ y $B = C$, entonces $A = C$ (propiedad transitiva).

Nótese que cualquier relación entre pares de objetos matemáticos que posee estas propiedades se llama relación de equivalencia. La igualdad de dos matrices significa que éstas de hecho son idénticas y, por tanto, una matriz puede ser sustituida por cualquier matriz igual en las operaciones sobre matrices.

### 2.2.2. Adición de matrices

**Definición 2.3.** *La **suma**, $A + B$, de dos matrices $A = [a_{ij}]$ y $B = [b_{ij}]$ del mismo orden $m \times n$ es la matriz $C = [c_{ij}]$ de $m \times n$ con elementos $c_{ij} = a_{ij} + b_{ij}$.*

Se dice que dos matrices del mismo orden son conformables para la suma. Puesto que la suma de dos matrices $m \times n$ cualesquiera de nuevo es una matriz $m \times n$, se dice que el conjunto de todas las matrices $m \times n$ es cerrado con respecto a la adición.

**Definición 2.4.** *Una matriz cuyos elementos son todos ceros se llama **matriz cero** y se denota por $O$ u $O_{(m,n)}$, cuando es necesario destacar el orden.*

La propiedad básica de la matriz $O_{(m,n)}$ es que para cualquier matriz $A$ de orden $m \times n$, se cumple la igualdad

$$A_{(m,n)} + O_{(m,n)} = A_{(m,n)}$$

pues la matriz cero es un elemento idéntico con respecto a la adición.

**Definición 2.5.** *La* ***negativa*** *o* ***negación*** *de una matriz $A$, denominada por $-A$, es la matriz del mismo orden que satisfaga la igualdad*

$$A + (-A) = O$$

*es decir, $-A$ es la matriz obtenida al multiplicar cada elemento de $A$ por $-1$.*

Dado que los elementos de las matrices son escalares, los cuales satisfacen la ley conmutativa y la asociativa de la adición, es fácil comprobar, basándonos en la definición de la suma de matrices, el siguiente teorema.

**Teorema 2.1.** *La adición de matrices es conmutativa y asociativa, esto es, si $A$, $B$ y $C$ son matrices del mismo orden, entonces*

$$A + B = B + A \tag{2.2}$$

$$A + (B + C) = (A + B) + C \equiv A + B + C \tag{2.3}$$

*y además*

$$A + O = A \tag{2.4}$$

$$A + (-A) = O \tag{2.5}$$

**Definición 2.6.** *El* ***producto*** *de una matriz $A = [a_{ij}]$ de orden $m \times n$ por un número (escalar) $k$, que se denota mediante $kA$, es la matriz $C = [c_{ij}]$ de orden $m \times n$ cuyos elementos son $c_{ij} = ka_{ij}$, es decir, que multiplicar una matriz por un número significa multiplicar todos los elementos de la matriz por este número.*

Basándose en las definiciones, se ve que para matrices cualesquiera de $m \times n$ y números cualesquiera, se tiene

$$k\,(A + B) = kA + kB \tag{2.6}$$

$$(k + q)\,A = kA + qA \tag{2.7}$$

$$k\,(qA) = (kq)\,A \tag{2.8}$$

$$1A = A \tag{2.9}$$

Nótese que $(-1)\,A = -A$, la negativa de $A$.

Las fórmulas, ecuaciones (2.2), (2.3), (2.4) y (2.5), así como el sistema de (2.6), (2.7), (2.8) y (2.9), determinan un espacio vectorial (véase capítulo 5).

**Teorema 2.2.** ***(Espacios vectoriales de matrices)*** *Todas las matrices reales (o complejas) del mismo orden $m \times n$ forman un espacio vectorial real (o complejo) de dimensión $mn$. Una base consta de todas las $mn$ matrices $E(r,c) = [e_{ij}(r,c)]$ $(r = 1,2,\ldots,m;\ c = 1,2,\ldots,n)$, en donde $e_{ij}(r,c) = \delta_{ir}\delta_{jc}$, pues el elemento en el $r$-ésimo renglón y en la $c$-ésima columna es igual a 1, $e_{rc}(r,c) = 1$, y los demás elementos son ceros.*

**Problema.** *Describir la base del espacio lineal de matrices de a) $3 \times 1$; b) $2 \times 2$ y c) $2 \times 3$.*

**Problema.** *Por definición, una matriz positiva es una matriz real cuadrada con todos sus elementos positivos. ¿Forman un espacio vectorial las matrices positivas de $m \times m$?*

### 2.2.3. Transpuesta y conjugada de una matriz. Matrices (anti)simétrica y (anti)hermitiana

En esta sección se definen otras operaciones sencillas sobre matrices y se presentan ciertos tipos de matrices especiales que tienen importancia en la práctica.

**Definición 2.7.** *La **transpuesta** $A^T$ de una matriz $A = [a_{ij}]$ de $m \times n$ es la matriz de $n \times m$ que se obtiene al intercambiar los renglones y las columnas en $A$, esto es, el $i$-ésimo renglón de $A$ se convierte en la $i$-ésima columna de $A^T$: $a_{ij}^T = a_{ji}$.*

**Definición 2.8.** *La **transpuesta** o **conjugada hermitiana** $A^+$ de una matriz $A = [a_{ij}]$ de $m \times n$ es la matriz de $n \times m$ que se obtiene al tomar el complejo conjugado de todos los elementos en la transpuesta $A^T$, esto es, $a_{ij}^+ = a_{ji}^*$.*

Por ser conmutativas las operaciones de transpuesta y de conjugación compleja, la transpuesta hermitiana de una matriz $A$ puede ser obtenida, primero, tomando la conjugada compleja $A^*$ de $A$ y, después, la transpuesta $(A^*)^T$, es decir, $A^+ = (A^T)^* = (A^*)^T$. Nótese que para las matrices reales la transpuesta y la transpuesta hermitiana son iguales.

Se queda de tarea para los estudiantes probar las siguientes propiedades:

$$(A^T)^T = A, \quad (A^+)^+ = A \tag{2.10}$$

$$(kA)^T = kA^T, \quad (kA)^+ = k^*A^+ \tag{2.11}$$

$$(A \pm B)^T = A^T \pm B^T, \quad (A \pm B)^+ = A^+ \pm B^+ \tag{2.12}$$

**Ejemplo 2.6.** *La transpuesta $A^T$ y conjugada hermitiana $A^+$ de una matriz $A$ son, respectivamente,*

$$A = \begin{bmatrix} -1 & 2i & 4 \\ 3 & 6 & i \end{bmatrix}, \quad A^T = \begin{bmatrix} -1 & 3 \\ 2i & 6 \\ 4 & i \end{bmatrix}, \quad A^+ = \begin{bmatrix} -1 & 3 \\ -2i & 6 \\ 4 & -i \end{bmatrix}$$

*De este ejemplo es fácil ver que la transpuesta de una matriz columna es una matriz renglón, y viceversa.*

**Definición 2.9.** *Se dice que una matriz $A = [a_{ij}]$ es **simétrica** si es igual a su transpuesta,*

$$A^T = A, \quad \text{es decir,} \quad a_{ji} = a_{ij} \tag{2.13}$$

*Se dice que una matriz $A = [a_{ij}]$ es **antisimétrica** si*

$$A^T = -A, \quad \text{es decir,} \quad a_{ji} = -a_{ij} \tag{2.14}$$

En el caso de una matriz $A$ antisimétrica, para $i = j$ se tiene $a_{ii} = -a_{ii}$, lo cual implica que todos los elementos que se encuentran en la diagonal principal de una matriz antisimétrica son ceros.

Se comprueba fácilmente que, dada cualquier matriz cuadrada $A$, la matriz $A + A^T$ es simétrica, mientras que $A - A^T$ es antisimétrica. De aquí sigue que cualquier matriz cuadrada $A$ puede escribirse de manera única como la suma de la matriz simétrica $\left(A + A^T\right)/2$ y la matriz antisimétrica $\left(A - A^T\right)/2$,

$$A = \left(A + A^T\right)/2 + \left(A - A^T\right)/2$$

**Definición 2.10.** *Se dice que una matriz $A = [a_{ij}]$ es **hermitiana** si es igual a su transpuesta hermitiana,*

$$A^+ = A, \quad \text{es decir,} \quad a_{ji}^* = a_{ij} \tag{2.15}$$

*Se dice que una matriz $A = [a_{ij}]$ es **antihermitiana** si*

$$A^+ = -A, \quad \text{es decir,} \quad a_{ji}^* = -a_{ij} \tag{2.16}$$

En el caso de una matriz $A$ hermitiana, para $i = j$ se tiene $a_{ii}^* = a_{ii}$, lo cual implica que todos los elementos que se encuentran en la diagonal principal de una matriz hermitiana son números reales. En el caso de una matriz $A$ antihermitiana, para $i = j$ se tiene $a_{ii}^* = -a_{ii}$, lo cual implica que todos los elementos que se encuentran en la diagonal principal de una matriz antihermitiana son números imaginarios o ceros.

Se comprueba fácilmente que, dada cualquier matriz cuadrada $A$, la matriz $A + A^+$ es hermitiana, mientras que $A - A^+$ es antihermitiana. De aquí se sigue que cualquier matriz cuadrada $A$ puede escribirse de manera única como la suma de la matriz hermitiana $(A + A^+)/2$ y la matriz antihermitiana $(A - A^+)/2$,

$$A = \left(A + A^+\right)/2 + \left(A - A^+\right)/2$$

**Ejemplo 2.7.** *Considere las matrices*

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix},$$

$$D = \begin{bmatrix} 2 & 1 + 2i \\ 1 - 2i & -5 \end{bmatrix}, \quad \bar{D} = \begin{bmatrix} 2 & 1 + 2i \\ 1 + 2i & -5 \end{bmatrix},$$

$$\tilde{D} = \begin{bmatrix} 2i & 1 + 2i \\ -1 + 2i & -5i \end{bmatrix}$$

*Entonces $A$ es simétrica y hermitiana, $B$ es antisimétrica y antihermitiana, $C$ no es (anti)simétrica ni (anti)hermitiana, $D$ es hermitiana pero no simétrica, $\bar{D}$ es simétrica pero no hermitiana y $\tilde{D}$ es antihermitiana.*

### 2.2.4. Multiplicación de matrices

En relación con las transformaciones lineales, que son transformaciones de un sistema de coordenadas a otro o mapeos entre espacios lineales, surge la operación de **multiplicación** de dos matrices. Como se verá en capítulo 5, esta operación representa dos mapeos consecutivos.

**Definición 2.11.** *Sea $A = [a_{ij}]$ una matriz de $m \times n$ y $B = [b_{jk}]$ una matriz de $n \times p$. Entonces el* **producto** *$AB$ (en este orden) se define como la matriz $C = [c_{ik}]$ de $m \times p$ cuyos elementos son*

$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk} \tag{2.17}$$

Nótese que para que el producto $AB$ esté definido, el número de columnas en $A$ debe ser igual al número de renglones en $B$. Se ve que el elemento $c_{ik}$ es el producto interior (producto punto) del $i$-ésimo vector renglón de la matriz $A$ por el $k$-ésimo vector columna de la matriz $B$.

**Propiedad 1.** Es fácil ver que, en general, la multiplicación de matrices no es conmutativa. Si $A$ y $B$ son matrices tales que tanto $AB$ como $BA$ están definidas, lo que implica que $A$ es de $m \times n$ y $B$ es de $n \times m$, entonces, en general,

$$AB \neq BA \tag{2.18}$$

Un ejemplo correspondiente de matrices cuadradas que comprueba esta propiedad es el siguiente:

$$\begin{aligned} A &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}, \\ AB &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \neq BA = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} \end{aligned}$$

Se dice que las matrices $A$ y $B$ son conmutativas si se cumple la igualdad $AB = BA$. La multiplicación de una matriz cuadrada $A$ por sí misma es conmutativa, $AA = AA$. Esto nos permite definir la **potencia entera positiva** $A^n$ de una matriz cuadrada $A$ de la manera natural:

$$A^1 = A,\ A^2 = AA,\ A^3 = AA^2 = AAA,\ \ldots,\ A^n = AAA\cdots A\ (n \text{ factores}) \tag{2.19}$$

De aquí se sigue que $A^p A^q = A^{p+q} = A^q A^p$.

**Propiedad 2.** Es fácil verificar que

$$A_{(m,n)} O_{(n,p)} = O_{(m,p)}, \quad O_{(m,n)} A_{(n,p)} = O_{(m,p)} \tag{2.20}$$

Sin embargo, al contrario que con los números, tiene lugar la siguiente propiedad:

**Propiedad 3.** En general, no se cumple la ley de cancelación, es decir,

$$AB = O \text{ no implica que } A = O \text{ o } B = O \tag{2.21}$$

Un ejemplo correspondiente es el siguiente:

$$\begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

**Propiedad 4.** La transposición de un producto es igual al producto de los factores transpuestos, tomados en orden inverso,

$$(AB)^T = B^T A^T \tag{2.22}$$

Para demostrar dicha propiedad suponemos que $A$ es de $m \times n$ y $B$ es de $n \times l$, de tal manera que la matriz $AB$ esté definida y sea de $m \times l$, quedando $(AB)^T$ de $l \times m$. Luego, $B^T$ es de $l \times n$, $A^T$ es de $n \times m$ y, por lo tanto, el producto $B^T A^T$ está definido. Para el $(ij)$-ésimo elemento de la matriz $(AB)^T$ se tiene

$$\left\langle (AB)^T \right\rangle_{ij} = \langle AB \rangle_{ji} = \sum_{q=1}^{n} a_{jq} b_{qi} = \sum_{q=1}^{n} a_{qj}^T b_{iq}^T = \sum_{q=1}^{n} b_{iq}^T a_{qj}^T = \langle B^T A^T \rangle_{ij}$$

lo que completa la demostración.

De manera similar se comprueba la siguiente propiedad.

**Propiedad 5.** La conjugada hermitiana de un producto de dos matrices es igual al producto de los factores transpuestos hermitianos, tomados en orden inverso,

$$(AB)^+ = B^+ A^+ \tag{2.23}$$

**Definición 2.12.** *Una matriz diagonal en la que todos los elementos de la diagonal principal son iguales, se llama **matriz escalar**,*

$$E = [e_{ij}] = \begin{bmatrix} k & 0 & \cdots & 0 \\ 0 & k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & k \end{bmatrix}$$

*en donde $e_{ij} = k\delta_{ij}$, $\delta_{ij}$ es el símbolo de Kronecker y $k$ es un número cualquiera.*

El nombre de la matriz $E$ se debe a la siguiente propiedad:

$$A_{(m,n)} E_{(n,n)} = E_{(m,m)} A_{(m,n)} = k A_{(m,n)} \tag{2.24}$$

es decir, la multiplicación de una matriz rectangular $A$ de $m \times n$ por una matriz escalar $E_{(n,n)}$ de la derecha o por una matriz escalar $E_{(m,m)}$ de la izquierda es igual a la multiplicación de $A$ por un escalar $k$. Además, de la ecuación (2.24) se obtiene que $E$ conmuta con cualquier matriz cuadrada del mismo orden. En el caso particular, $k = 1$, de la matriz escalar $E$ se tiene la **matriz identidad** o **unidad** $I$, tal que

$$AI = IA = A \tag{2.25}$$

para cualquier matriz cuadrada $A$ del mismo orden. Por definición $A^0 = I$.

**Teorema 2.3.** ***(Propiedades de la multiplicación de matrices)*** *La multiplicación de matrices es*

*1) asociativa*

$$A(BC) = (AB)C \tag{2.26}$$

$$(kA)B = k(AB) = A(kB) \tag{2.27}$$

*2) distributiva con respecto a la adición de matrices*

$$A(B+C) = AB+AC \tag{2.28}$$

$$(A+B)C = AC+BC \tag{2.29}$$

**Demostración.** Supongamos que $A$ es de $n\times l$, $B$ es de $l\times m$ y $C$ es de $m\times p$. Entonces para el $(ij)$-ésimo elemento de la matriz $A(BC)$ se tiene

$$\begin{aligned}\langle A(BC)\rangle_{ij} &= \sum_{q=1}^{l} a_{iq}\langle BC\rangle_{qj} = \sum_{q=1}^{l} a_{iq}\left(\sum_{s=1}^{m} b_{qs}c_{sj}\right)\\ &= \sum_{q=1}^{l}\sum_{s=1}^{m} a_{iq}b_{qs}c_{sj} = \sum_{s=1}^{m}\left(\sum_{q=1}^{l} a_{iq}b_{qs}\right)c_{sj}\\ &= \sum_{s=1}^{m}\langle AB\rangle_{is}c_{sj} = \langle (AB)C\rangle_{ij}\end{aligned}$$

que completa la demostración de la ecuación (2.26). De manera similar o usando propiedades de la matriz escalar $E=kI$, se comprueba la ecuación (2.27), que se deja como ejercicio. La propiedad distributiva, ecuaciones (2.28) y (2.29), se demuestra fácilmente, y también se deja como ejercicio.

**Problema.** *Demuestre las igualdades* $(ABC)^T = C^TB^TA^T$ *y* $(ABC)^+ = C^+B^+A^+$; $ABCD = A(B(CD)) = (AB)(CD) = A((BC)D) = (A(BC))D$.

**Problema.** *¿En caso de que* $AB = AC$ *o* $BA = CA$ *se sigue necesariamente que* $B = C$*?*

**Problema.** *¿Cuándo es válida la igualdad* $(A+B)(A-B) = A^2 - B^2$*?*

## 2.3. Matrices inversas

### 2.3.1. Definición y propiedades

Dado el número $a$, su inverso recíproco $x$ es el número tal que resuelva la ecuación $xa = ax = 1$. Por cierta analogía con los números y tomando en cuenta que, en general, el producto de matrices no es conmutativo, podemos introducir el concepto de la matriz inversa $X$ de una matriz dada $A$, de tal manera que la inversa resuelva $XA = I$ y/o $AX = I$.

**Definición 2.13.** *Sea $A$ una matriz rectangular de $m \times n$.*

*a) Cualquier matriz $L_{(n,m)}$ para la cual $L_{(n,m)}A_{(m,n)} = I_{(n,n)}$ se llama **inversa izquierda** de $A$.*

*b) Cualquier matriz $R_{(n,m)}$ para la cual $A_{(m,n)}R_{(n,m)} = I_{(m,m)}$ se llama **inversa derecha** de $A$.*

*c) Cualquier matriz $A^{-1}_{(n,m)}$ para la cual $A^{-1}_{(n,m)}A_{(m,n)} = I_{(n,n)}$ y $A_{(m,n)}A^{-1}_{(n,m)} = I_{(m,m)}$ se llama **inversa (bilateral)** de $A$.*

Si la matriz $A$ es de $m \times n$, entonces cualquier inversa derecha $R$ o izquierda $L$ de $A$ y, por lo tanto, la inversa bilateral $A^{-1}$ deben ser de $n \times m$. Se puede determinar cuál de estos tipos de inversas tiene una matriz $A$ mediante la solución de conjuntos de ecuaciones lineales correspondientes, en donde los elementos de las inversas están considerados como incógnitas. En caso de que $A$ sea de $n \times n$, las inversas, si existen, son matrices cuadradas del mismo orden. Debe mencionarse que un análisis más a fondo de la existencia de inversas bilaterales nos llevará a la conclusión de que las inversas bilaterales existen sólo para las matrices cuadradas, tanto que en el inciso $c$ de la definición anterior hay que suponer $m = n$.

**Ejemplo 2.8.** *Buscaremos una matriz inversa derecha $R$ para la matriz*

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}$$

*que requiere resolver la ecuación $AR = I$, esto es,*

$$\begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x & z \\ y & w \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

*Multiplicando las matrices se tienen dos conjuntos de ecuaciones lineales:*

$$\begin{array}{ll} x - y = 1 & z - w = 0 \\ x + 2y = 0 & z + 2w = 1 \end{array}$$

*La solución de este sistema de ecuaciones es única: $x = 2/3$, $y = -1/3$, $z = 1/3$, $w = 1/3$. Por lo tanto, hay exactamente una matriz inversa derecha, que es*

$$R = \begin{bmatrix} 2/3 & 1/3 \\ -1/3 & 1/3 \end{bmatrix}$$

**Ejemplo 2.9.** *Se busca una matriz inversa derecha $R$ para la matriz*

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 2 \end{bmatrix}.$$

*que implica la resolución de la ecuación $AR = I$, esto es,*

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} x & u \\ y & v \\ z & w \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

*La última se reduce al sistema de ecuaciones lineales subdeterminadas, es decir, el número de incógnitas es mayor que el número de ecuaciones:*

$$x-y+z=1 \qquad u-v+w=0$$
$$x+y+2z=0 \qquad u+v+2w=1$$

*Al resolverlas, se pueden dar valores arbitrarios para* $z=\alpha$ *y* $w=\beta$. *Resolviendo con respecto a las otras variables y considerando* $\alpha$ *y* $\beta$ *como parámetros, finalmente se obtiene un número infinito de inversas derechas*

$$R=\begin{bmatrix} (1-3\alpha)/2 & (1-3\beta)/2 \\ -(1+\alpha)/2 & (1-\beta)/2 \\ \alpha & \beta \end{bmatrix}$$

**Ejemplo 2.10.** *Se busca una inversa izquierda* $L$ *de la matriz*

$$A=\begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix} \quad \text{en la forma} \quad L=\begin{bmatrix} x & y \\ z & w \end{bmatrix}$$

*De la ecuación* $LA=I$ *se obtiene un sistema inconsistente de ecuaciones lineales,*

$$x-2y=1 \qquad z-2w=0$$
$$-x+2y=0 \qquad -z+2w=1$$

*que no tiene solución. Por lo tanto, no existe la inversa izquierda para la matriz* $A$ *dada.*

Los ejemplos anteriores muestran la manera de encontrar matrices inversas y, además, el hecho de que las inversas pueden no existir o, al contrario, ser múltiples. El siguiente teorema aclara la existencia y unicidad de inversas bilaterales.

**Teorema 2.4.** ***(Inversas bilaterales)*** *Sea* $A$ *una matriz. Si existen una inversa derecha* $R$ *y una inversa izquierda* $L$ *de* $A$, *entonces son iguales y son una inversa bilateral* $A^{-1}$. *Cualesquiera dos inversas bilaterales de* $A$ *son idénticas.*

**Demostración.** *Sean* $R$ *y* $L$ *inversas derecha e izquierda de* $A$, *respectivamente. Entonces,*

$$R=IR=(LA)R=L(AR)=LI=L$$

*de modo que* $R=L\equiv A^{-1}$, *y ésta es una inversa bilateral. Luego, supongamos que* $X$ *y* $Y$ *son dos inversas bilaterales de* $A$. *A partir de*

$$X=IX=(YA)X=Y(AX)=YI=Y$$

*se obtiene la unicidad.* ■

La propiedad de tener la inversa bilateral es tan importante que las matrices cuadradas con tales inversas se distinguen con un nombre especial.

**Definición 2.14.** *Si una matriz $A$, necesariamente cuadrada, tiene la inversa bilateral $A^{-1}$, $AA^{-1} = A^{-1}A = I$, se dice que $A$ es una* **matriz no singular***. Una* **matriz singular** *es una matriz cuadrada que no tiene inversa bilateral.*

**Teorema 2.5.** *Sean $A$ y $B$ matrices cuadradas no singulares de $n \times n$. Entonces,*

*a) $AB$ es no singular y $(AB)^{-1} = B^{-1}A^{-1}$;*

*b) $A^{-1}$ es no singular y $\left(A^{-1}\right)^{-1} = A$;*

*c) $A^T$ y $A^+$ son no singulares con $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$ y $\left(A^+\right)^{-1} = \left(A^{-1}\right)^+$.*

**Demostración.** *La matriz $B^{-1}A^{-1}$ es de $n \times n$. De $\left(B^{-1}A^{-1}\right)(AB) = B^{-1}\left(A^{-1}A\right)B = B^{-1}IB = I$ se tiene que $B^{-1}A^{-1}$ es la inversa izquierda de $A$. A la vez, de $(AB)\left(B^{-1}A^{-1}\right) = A\left(BB^{-1}\right)A^{-1} = AIA^{-1} = I$ se tiene que $B^{-1}A^{-1}$ es la inversa derecha de $A$. Entonces, según el teorema de inversas bilaterales, la matriz $AB$ tiene la inversa bilateral única, $(AB)^{-1} = B^{-1}A^{-1}$, y por lo tanto es no singular. Luego, como $A$ es no singular, $A^{-1}$ existe por definición. Es fácil ver que $A$ es inversa bilateral de $A^{-1}$. Entonces, $\left(A^{-1}\right)^{-1} = A$ y la matriz $A^{-1}$ es no singular. De las igualdades $I = I^T = \left(AA^{-1}\right)^T = \left(A^{-1}\right)^T A^T$ y $I = I^T = \left(A^{-1}A\right)^T = A^T\left(A^{-1}\right)^T$ se sigue que $\left(A^{-1}\right)^T$ es la inversa bilateral de $A^T$, esto es, $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$ y $A^T$ es no singular. De manera similar se comprueba que $A^+$ es no singular con $\left(A^+\right)^{-1} = \left(A^{-1}\right)^+$.* ■

### 2.3.2. Matrices inversas y sistemas de ecuaciones lineales

Un sistema de $m$ ecuaciones lineales con $n$ incógnitas $x_1, x_2, \ldots, x_n$ en forma matricial es

$$A\mathbf{x} = \mathbf{b} \tag{2.30}$$

en donde $A = [a_{ij}]$ es una matriz conocida de $m \times n$, $\mathbf{x}$ es un vector columna (matriz) de $n \times 1$ compuesto de $n$ incógnitas $x_1, x_2, \ldots, x_n$, $\mathbf{b}$ es un vector columna (matriz) de $m \times 1$ de componentes conocidos $b_1, b_2, \ldots, b_m$. La apariencia formal de la ecuación matricial y su homólogo de números, $ax = b$, nos sugiere resolver la ecuación (2.30) mediante la matriz inversa. Dado que las matrices inversas no siempre existen o una matriz puede tener muchas inversas, la situación no es tan simple como en caso de números. Se conoce que para un sistema de ecuaciones lineales puede cumplirse una de las tres posibilidades: no existe ninguna solución, existe solución única y existe un número infinito de soluciones. La relación de la solución de los sistemas de ecuaciones lineales con varias matrices inversas se aclara mediante el siguiente teorema.

**Teorema 2.6.** *Sea $A\mathbf{x} = \mathbf{b}$ la representación de un sistema de $m$ ecuaciones lineales con $n$ incógnitas.*

*a) Suponemos que $A$ tiene no más de una inversa izquierda $L$. Si existen soluciones $\mathbf{x}$ a la ecuación $A\mathbf{x} = \mathbf{b}$, entonces deben ser iguales a $L\mathbf{b}$. Por lo tanto, existe cuando mucho una solución, que es $L\mathbf{b}$.*

*b) Suponemos que $A$ tiene una inversa derecha $R$. Entonces, existe al menos una solución, que es $R\mathbf{b}$, a la ecuación $A\mathbf{x} = \mathbf{b}$.*

*c) Suponemos que $A$ es no singular. Entonces existe exactamente una solución a $A\mathbf{x} = \mathbf{b}$, es decir, $A^{-1}\mathbf{b}$.*

**Demostración.** *a) Si $\mathbf{x}_0$ es alguna solución a $A\mathbf{x} = \mathbf{b}$, entonces $\mathbf{b} = A\mathbf{x}_0$. Multiplicando la última ecuación por $L$ se tiene*

$$L\mathbf{b} = L\,(A\mathbf{x}_0) = (LA)\,\mathbf{x}_0 = I\mathbf{x}_0 = \mathbf{x}_0$$

*Pues, si existe una solución, ésta debe ser igual a $L\mathbf{b}$.*

*b) Se tiene que*

$$A\,(R\mathbf{b}) = (AR)\,\mathbf{b} = I\mathbf{b} = \mathbf{b}$$

*por lo tanto, $R\mathbf{b}$ es una solución.*

*c) Si $A$ es no singular, existe la inversa bilateral $A^{-1}$. Ya que $A^{-1}$ es una inversa derecha, $A^{-1}\mathbf{b}$ es una solución según el inciso anterior. Por otra parte, cualquier solución debe ser igual a $A^{-1}\mathbf{b}$ según el inciso a, ya que $A^{-1}$ es una inversa izquierda. Además, la matriz $A^{-1}$ es única. Entonces, la solución $A^{-1}\mathbf{b}$ es única también. Esto completa la demostración de la existencia y unicidad de la solución.* ■

De lo anterior se ve cómo se pueden solucionar sistemas de ecuaciones lineales mediante matrices inversas. Pero resulta que cada matriz inversa a la vez se obtiene mediante la solución de sistemas de ecuaciones lineales. Pues, volvemos exactamente a problemas del mismo tipo que se querían resolver desde un principio. En la práctica, por lo general, es más eficiente resolver un sistema de ecuaciones lineales directamente (por ejemplo, usando el método de eliminación de Gauss o de Gauss-Jordan), que calcular primero $A^{-1}$ y luego $A^{-1}\mathbf{b}$. Sin embargo, el concepto y representación de la solución como $\mathbf{x} = A^{-1}\mathbf{b}$ pueden ser muy poderosos para el entendimiento general de ciertos aspectos en muchos problemas. Proposiciones generales acerca de problemas de existencia y unicidad de solución de un sistema de ecuaciones lineales se dan mediante el concepto de rango de una matriz, y la información relevante se puede encontrar en varios libros, por ejemplo, véase [8], [11] y [16].

## 2.4. Matriz unitaria, ortogonal. Traza

Las matrices unitarias y ortogonales están estrechamente relacionadas con transformaciones entre las bases ortonormales en espacios vectoriales lineales (véase el capítulo 5). Dada la importancia de dichos conceptos, en esta sección se presenta la definición y algunas de las propiedades importantes de las matrices unitarias.

Sean $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ y $\{\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n\}$ dos bases en un espacio vectorial lineal $L$. Sea $T$ la matriz de una transformación de coordenadas, $x' = Tx$ de un vector $\mathbf{x}$ que tiene coordenadas $x'$ en la base $\{\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n\}$ y coordenadas $x$ en la base $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$. Además, sea $A$ la matriz que representa un mapeo

lineal, $y = Ax$, del espacio $L$ en sí considerado en la base $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$. Se sabe de la teoría de espacios vectoriales lineales que la matriz $A'$ de dicho mapeo lineal en la base $\{\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n\}$ se describe mediante la **transformación de semejanza** $A' = TAT^{-1}$. Dos matrices cualesquiera $A'$ y $A$ relacionadas mediante la transformación de semejanza $A' = TAT^{-1}$, donde $T$ es una matriz no singular, se llaman semejantes.

**Definición 2.15.** *Una matriz $A$, tal que $A^+ = A^{-1}$, se llama **matriz unitaria**. Una **matriz ortogonal** es una matriz unitaria real, tal que $A^T = A^{-1}$.*

La clase de matrices unitarias incluye la clase de matrices ortogonales. Por lo tanto, cualquier afirmación válida para toda matriz unitaria es también válida para toda matriz ortogonal.

Antes de proceder con las propiedades de matrices unitarias introducimos el concepto de traza de una matriz, que es muy útil en varios problemas de la teoría de matrices.

**Definición 2.16.** *La **traza de una matriz** cuadrada $A$ se define como la suma de los elementos en la diagonal principal de $A$, es decir,*

$$trA = \sum_{i=1}^{n} a_{ii} \tag{2.31}$$

La propiedad importante de la traza es que

$$trAB = trBA \tag{2.32}$$

La demostración es simple:

$$\begin{aligned} trAB &= \sum_{i=1}^{n} \langle AB \rangle_{ii} = \sum_{i=1}^{n} \left( \sum_{j=1}^{n} a_{ij} b_{ji} \right) = \sum_{j=1}^{n} \left( \sum_{i=1}^{n} b_{ji} a_{ij} \right) = \sum_{j=1}^{n} \langle BA \rangle_{jj} \\ &= trBA \end{aligned}$$

De la propiedad anterior se tiene que, si una matriz $S$ es no singular, entonces

$$tr\left(S^{-1}AS\right) = tr\left(ASS^{-1}\right) = trA \tag{2.33}$$

Las matrices unitarias tienen propiedades importantes.

**Propiedad 1.** Una matriz cuadrada $A = [a_{ij}]$ es unitaria (u ortogonal) si y sólo si sus vectores columna (y también los vectores renglón) forman un sistema de vectores ortonormales.

**Demostración.** *Suponemos que $A$ es una matriz unitaria con vectores columna $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$, es decir, $A = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n]$. Su conjugada hermitiana $A^+ = A^{-1}$ y, por lo tanto,*

$$I = A^{-1}A = A^+A = \begin{bmatrix} \mathbf{a}_1^+ \\ \mathbf{a}_2^+ \\ \vdots \\ \mathbf{a}_n^+ \end{bmatrix} [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n] = \begin{bmatrix} \mathbf{a}_1^+\mathbf{a}_1 & \mathbf{a}_1^+\mathbf{a}_2 & \cdots & \mathbf{a}_1^+\mathbf{a}_n \\ \mathbf{a}_2^+\mathbf{a}_1 & \mathbf{a}_2^+\mathbf{a}_2 & \cdots & \mathbf{a}_2^+\mathbf{a}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_n^+\mathbf{a}_1 & \mathbf{a}_n^+\mathbf{a}_2 & \ldots & \mathbf{a}_n^+\mathbf{a}_n \end{bmatrix}$$

*en donde* $\mathbf{a}_i^+\mathbf{a}_j$ *es el producto interior (producto punto) de dos vectores* $\mathbf{a}_i$ *y* $\mathbf{a}_j$. *De la última fórmula se tiene que*

$$\mathbf{a}_i^+\mathbf{a}_j = \delta_{ij}, \quad (\delta_{ij} \text{ es la delta de Kronecker, } i,j = 1,2,\ldots,n)$$

*que demuestra la ortonormalidad del conjunto de vectores columna* $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$. *De manera similar se comprueba la ortonormalidad del conjunto de vectores renglón de una matriz unitaria.*

*Suponemos que los vectores columna* $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$, *de una matriz* $A$ *son ortonormales,* $\mathbf{a}_i^+\mathbf{a}_j = \delta_{ij}$. *Esto implica que* $A^+A = I$. *Por lo tanto,* $A^+$ *es una inversa izquierda de* $A$. *Por otra parte, los vectores* $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$, *siendo ortonormales por suposición, son linealmente independientes y, por lo tanto, el rango de la matriz* $A$ *es igual a* $n$. *Esto significa que* $A$ *es no singular y existe su inversa bilateral* $A^{-1}$. *Multiplicando* $A^+A = I$ *por* $A^{-1}$ *por la derecha, se tiene* $A^+ = A^{-1}$, *es decir,* $A$ *es una matriz unitaria. De manera similar se comprueba que una matriz con vectores renglón ortonormales es una matriz unitaria.* ∎

**Propiedad 2.** Si $A$ es una matriz unitaria, entonces $A^{-1}$ es unitaria también, es decir, $\left(A^{-1}\right)^+ = \left(A^{-1}\right)^{-1} = A$.

**Demostración.** *Por definición,* $A^{-1} = A^+$. *Luego,* $\left(A^{-1}\right)^+ = \left(A^+\right)^+ = A$. *Entonces, multiplicando la última igualdad por* $A^{-1}$ *por la derecha e izquierda, respectivamente, se tiene que* $\left(A^{-1}\right)^+ A^{-1} = AA^{-1} = I$ *y* $A^{-1}\left(A^{-1}\right)^+ = A^{-1}A = I$, *es decir,* $\left(A^{-1}\right)^+ = \left(A^{-1}\right)^{-1} = A$. ∎

**Propiedad 3.** El producto $AB$ de dos matrices $A$ y $B$ unitarias también es una matriz unitaria, es decir, $(AB)^+ = (AB)^{-1} = B^{-1}A^{-1}$.

**Demostración.** *La demostración es como sigue:*

$$\begin{aligned}
(AB)^+ &= B^+A^+ = B^{-1}A^{-1} \\
(AB)^+(AB) &= \left(B^{-1}A^{-1}\right)(AB) = B^{-1}\left(A^{-1}A\right)B = I \\
(AB)(AB)^+ &= (AB)\left(B^{-1}A^{-1}\right) = A\left(BB^{-1}\right)A^{-1} = I
\end{aligned}$$

*De las dos últimas líneas se sigue que* $(AB)^+$ *es la inversa bilateral de* $AB$. *Esto concluye la demostración.* ∎

**Propiedad 4.** Si $A$ es una matriz unitaria y si $(\mathbf{x}, \mathbf{y}) = \mathbf{x}^+\mathbf{y}$ es el producto interior estándar de dos vectores columna, entonces, $(A\mathbf{x}, A\mathbf{y}) = (\mathbf{x}, \mathbf{y})$. Lo último sigue de $(A\mathbf{x}, A\mathbf{y}) = (A\mathbf{x})^+(A\mathbf{y}) = \mathbf{x}^+\left(A^+A\right)\mathbf{y} = \mathbf{x}^+\mathbf{y} = (\mathbf{x}, \mathbf{y})$.

## 2.5. Determinante e inversa de una matriz no singular

El determinante de una matriz cuadrada $A$ de orden $n$ es un escalar asociado con esta matriz. El concepto de determinante surge de manera natural en el proceso de resolución de un sistema de $n$ ecuaciones lineales con $n$ incógnitas

y, por tanto, los determinantes tienen diversas aplicaciones en las Matemáticas, principalmente en las que se relacionan con las soluciones de un sistema de ecuaciones lineales. Conservando su importancia en problemas de matemáticas teóricas (véase, por ejemplo, la ecuación característica), resulta que son poco prácticos para cálculos numéricos aún en el problema de la resolución de sistemas de $n$ ecuaciones lineales con $n$ incógnitas, debido a que no conducen a una solución en el caso de una matriz cuadrada singular y, aún menos, en casos de matrices rectangulares. Anotamos que el método general de solución de un sistema cualquiera de ecuaciones lineales se proporciona mediante la matriz pseudoinversa (el método de mínimos cuadrados).

El determinante puede definirse de varias maneras equivalentes. Basándonos en la experiencia con las matrices cuadradas de segundo y tercer orden, podemos ver que la manera más práctica de hacerlo es por inducción.

**Definición 2.17.** *Sea $A$ una matriz cuadrada de $n \times n$. El* ***menor*** *$M_{ij}$ de $A$ es el determinante de la matriz de $(n-1) \times (n-1)$ formada de los elementos de $A$ al omitir el i-ésimo renglón y la j-ésima columna de $A$. El* ***cofactor*** *$C_{ij}$ de $A$ es $C_{ij} = (-1)^{i+j} M_{ij}$.*

**Definición 2.18.** *El determinante* $\det A$ *de una matriz $A = [a]$ de $1 \times 1$ se define como* $\det A = a$. *El* ***determinante*** $\det A$ *de una matriz $A = [a_{ij}]$ de $n \times n$ se define como*

$$\det A = \sum_{j=1}^{n} a_{1j} C_{1j} \tag{2.34}$$

*es decir, el determinante de $A$ es la suma de los productos de los elementos del primer renglón y de los cofactores correspondientes.*

**Ejemplo 2.11.** *Para la matriz $A$ de $2 \times 2$,*

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

*se tiene*

$$\begin{aligned} \det A &= a_{11}C_{11} + a_{12}C_{12} \\ &= a_{11}\det[a_{22}] - a_{12}\det[a_{21}] \\ &= a_{11}a_{22} - a_{12}a_{21} \end{aligned}$$

Resulta que el $\det A$ es la suma de todos los productos posibles de $n$ elementos de $A$, con los signos adecuados, y donde en cada producto hay exactamente un término de cada renglón y cada columna. De hecho, los determinantes pueden definirse de este modo.

El primer renglón de $A$ jugó un papel especial en la definición del $\det A$. Sin embargo, la propiedad importante del determinante es que el primer renglón puede sustituirse por cualquier renglón o columna. Esta propiedad se da por el siguiente teorema, omitiendo la demostración por ser más complicada que instructiva.

**Teorema 2.7.** *Sea $A$ una matriz cuadrada de $n \times n$. El $\det A$ es igual a la suma de los productos de los elementos de cualquier renglón o de cualquier columna por sus cofactores correspondientes:*

$$\det A = \sum_{j=1}^{n} a_{1j} C_{1j} = \sum_{j=1}^{n} a_{rj} C_{rj} = \sum_{i=1}^{n} a_{ik} C_{ik} \tag{2.35}$$

Los propiedades importantes de los determinantes son las siguientes.

**Corolario.** *a) El determinante de una matriz y de su transpuesta son iguales*

$$\det A = \det A^T$$

*b) Si cualquier renglón o columna de $A$ se multiplica por un número c, quedando como resultado una $\tilde{A}$, entonces*

$$\det \tilde{A} = c \det A$$

*En particular, el determinante de una matriz con cualquier renglón o columna igual a cero es cero.*

*c) Para cualquier número c y una matriz $A$ de $n \times n$, se tiene*

$$\det cA = c^n \det A$$

*d) Si $A$ tiene dos renglones (columnas) iguales, entonces $\det A = 0$. De donde, con base en el inciso b se tiene que, si $A$ tiene dos renglones (columnas) múltiplos, entonces $\det A = 0$.*

*e) Si $A'$ se obtiene de $A$ intercambiando dos renglones (o dos columnas), entonces*

$$\det A' = (-1) \det A$$

*f) Si $A'$ se obtiene de $A$ sumando a un renglón (una columna) un múltiplo de otro renglón (otra columna), entonces*

$$\det A' = \det A$$

**Demostración.** *a) Desarrollar el $\det A$ con respecto a su primer renglón es lo mismo que desarrollar el $\det A'$ con respecto a su primera columna.*

*b) Desarrollando el $\det \tilde{A}$ con respecto al renglón (a la columna) que fue multiplicado por un número c, se tiene $\det \tilde{A} = c \det A$.*

*c) Se comprueba fácilmente por inducción.*

*d) El enunciado es válido para cualquier matriz $A$ de $2 \times 2$ con dos renglones iguales. Suponemos que es también válido para cualquier matriz $A$ de $(n-1) \times (n-1)$. Al desarrollar una matriz $A$ de $n \times n$ con respecto a un renglón que es diferente de los dos iguales se tiene que todos los productos son ceros según la suposición hecha con respecto a las matrices de $(n-1) \times (n-1)$. Entonces, el resultado es cierto por inducción.*

*e) La demostración es por inducción. El enunciado se cumple para determinantes de orden $n = 2$. Suponemos que es también válido para determinantes de orden $(n-1)$. Para cualquier matriz $A$ de orden $n$, supóngase que $A'$ se obtiene a partir de $A$ al intercambiar dos renglones. Desarrollando* $\det A'$ *y* $\det A$ *con respecto a un renglón $k$-ésimo que no sea uno de los dos intercambiados, se tiene*

$$\det A' = \sum_{j=1}^{n} a_{kj}\,(-1)^{k+j}\,M'_{kj}, \quad \det A = \sum_{j=1}^{n} a_{kj}(-1)^{k+j}\,M_{kj}$$

*La matriz correspondiente al menor $M'_{kj}$ de matriz $A'$ tiene dos renglones intercambiados con respecto a la matriz correspondiente al menor $M_{kj}$ de la matriz $A$. Por suposición, $M'_{kj} = (-1)\,M_{kj}$. Entonces,*

$$\det A' = \sum_{j=1}^{n} a_{kj}\,(-1)^{k+j}\,M'_{kj} = \sum_{j=1}^{n} a_{kj}\,(-1)^{k+j}\,(-1)\,M_{kj} = (-1)\det A$$

*f) Sea $A = [\mathbf{a}_1, \ldots, \mathbf{a}_k, \ldots, \mathbf{a}_n]$ una matriz cuadrada y sean $B = [\mathbf{a}_1, \ldots, \mathbf{b}_k, \ldots, \mathbf{a}_n]$ y $C = [\mathbf{a}_1, \ldots, \mathbf{a}_k + \mathbf{b}_k, \ldots, \mathbf{a}_n]$ dos matrices que se distinguen de $A$ en una sola columna $k$-ésima del modo indicado. Desarrollando* $\det C$ *con respecto a la columna $k$-ésima, se tiene que*

$$\det C = \det A + \det B$$

*En caso de que $\mathbf{b}_k = c\mathbf{a}_m$ $(m \neq k)$ se tiene*

$$\det A' = \det\,[\mathbf{a}_1, \ldots, \mathbf{a}_k + c\mathbf{a}_m, \ldots, \mathbf{a}_n] = \det A$$

*debido a que* $\det B = 0$, *según el inciso d.* ■

A continuación será enunciado sin demostración un teorema importante sobre el producto de dos matrices cuadradas. (La demostración se puede ver en [11] o [16])

**Teorema 2.8.** *Sean $A = [a_{ij}]$ y $B = [b_{ij}]$ matrices de $n \times n$. Entonces,*

$$\det AB = \det A \det B \tag{2.36}$$

Este teorema nos proporciona la base para probar el teorema sobre la inversa de una matriz cuadrada no singular.

**Teorema 2.9.** *Una matriz $A = [a_{ij}]$ es no singular si y sólo si* $\det A \neq 0$. *Si $A$ es no singular, entonces*

$$\det A^{-1} = (\det A)^{-1} \tag{2.37}$$

**Demostración.** *Nótese que la segunda parte del teorema se comprueba fácilmente con base en la ecuación (2.36). Si $A$ es no singular, entonces, existe su inversa $A^{-1}$ tal que $A^{-1}A = I$. De la ecuación (2.36) se tiene que*

$\det A^{-1}A = \det A^{-1} \det A = \det I = 1$, *de donde se sigue que tanto* $\det A \neq 0$ *como* $\det A^{-1} \neq 0$ *y se cumple la ecuación (2.37).*

*Si* $\det A \neq 0$, *la existencia de la inversa* $A^{-1}$ *se demuestra de modo constructivo. Demostramos que la inversa de* $A$ *es*

$$A^{-1} = \frac{C^T}{\det A}, \quad C = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{bmatrix} \tag{2.38}$$

*en donde los elementos de la matriz* $C$ *son cofactores de los elementos correspondientes de la matriz* $A$. *Denotamos* $B = C^T / \det A$, *y sea* $BA = D = [d_{kl}]$. *Por definición de la multiplicación de matrices*

$$d_{kl} = \frac{1}{\det A} \sum_{i=1}^{n} C_{ik} a_{il} = \begin{cases} 1 & para \quad k = l \\ 0 & para \quad k \neq l \end{cases}$$

*La última igualdad se debe a que* $\sum_{i=1}^{n} C_{ik} a_{ik}$ *es nada más que el desarrollo de* $\det A$ *con respecto a la* $k$*-ésima columna de* $A$. *Para* $k \neq l$, *éste es un desarrollo del determinante de la matriz* $\tilde{A}$ *que se obtiene de* $A$ *al reemplazar la* $k$*-ésima columna por la* $l$*-ésima, de modo que* $\tilde{A}$ *tiene dos columnas iguales y su determinante es cero. De donde,* $BA = D = I$. *De manera semejante,* $AB = I$. *Finalmente, se tiene que* $B = A^{-1}$. ■

## 2.6. Eigenvalores y eigenvectores

Los problemas de eigenvalores (valores propios) y eigenvectores de matrices y de transformaciones lineales son de importancia fundamental en las matemáticas aplicadas. Una de las razones es que esos conceptos básicos surgen en una gran variedad de modelos matemáticos de fenómenos físicos, de procesos de ingeniería y química. Cabe mencionar que los conceptos de espacio lineal y de eigenvalores y eigenvectores de operadores de observables físicos son fundamentales en la mecánica cuántica. En esta sección se consideran los conceptos básicos para los problemas de eigenvalores.

Sea $A = [a_{ij}]$ una matriz de $n \times n$ y considérese la ecuación vectorial

$$A\mathbf{x} = \lambda \mathbf{x} \tag{2.39}$$

en donde $\lambda$ es un número, real o complejo. Es evidente que el vector $\mathbf{x} = \mathbf{0}$ es una solución de esta ecuación para cualquier valor de $\lambda$. Un número $\lambda$ para el que la ecuación (2.39) tiene una solución no trivial, $\mathbf{x} \neq \mathbf{0}$, se llama **eigenvalor** (**valor propio** o **característico**) de la matriz $A$. Las soluciones $\mathbf{x} \neq \mathbf{0}$ de la ecuación (2.39) se llaman **eigenvectores** (**vectores propios** o **característicos**) de $A$ correspondientes a ese eigenvalor $\lambda$. El conjunto de los eigenvalores se llama **espectro** de $A$. El mayor valor absoluto de los eigenvalores de $A$ se llama el **radio espectral** de $A$. El problema de determinar los eigenvalores y eigenvectores de una matriz se conoce como **problema de eigenvalores**.

**Ejemplo 2.12.** *Para resolver la ecuación diferencial matricial $\ddot{\mathbf{z}} + A\mathbf{z} = \mathbf{0}$, se hace la sustitución $\mathbf{z} = \mathbf{x}e^{i\omega t}$ que conduce a $\omega^2 \mathbf{x}e^{i\omega t} = A\mathbf{x}e^{i\omega t}$, o bien, $A\mathbf{x} = \lambda\mathbf{x}$. Entonces, para que $\mathbf{z} = \mathbf{x}e^{i\omega t}$ sea la solución no trivial, el número $\lambda = \omega^2$ debe ser un eigenvalor de $A$ y el vector $\mathbf{x}$ debe ser un eigenvector correspondiente.*

**Ejemplo 2.13.** *Dada una transformación lineal $\mathbf{y} = A\mathbf{x}$, definida por una matriz $A$, se buscan vectores $\mathbf{x}$ tales que vectores correspondientes $\mathbf{y}$ tendrían la misma dirección y sentido que los de $\mathbf{x}$. Esto significa que se desea hallar vectores $\mathbf{x}$ tales que $\mathbf{y} = A\mathbf{x} = \lambda\mathbf{x}$. Estos vectores existen si y sólo si $A$ tiene eigenvalores reales positivos. Entonces, $\mathbf{x}$ son eigenvectores correspondientes a esos eigenvalores.*

Si $\mathbf{x}$ es un eigenvector, entonces $\mathbf{x}$ y $A\mathbf{x}$ son linealmente dependientes, es decir, los vectores $\mathbf{x}$ y $A\mathbf{x}$ son proporcionales y el factor de proporcionalidad es el eigenvalor $\lambda$. La ecuación (2.39) escrita en forma

$$(A - \lambda I)\,\mathbf{x} = \mathbf{0} \tag{2.40}$$

es un sistema homogéneo de ecuaciones lineales con respecto a componentes del vector $\mathbf{x}$. Por el teorema de Cramer, este sistema tiene por lo menos una solución no trivial si y sólo si el determinante del sistema es cero:

$$D(\lambda) \equiv \det(A - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0 \tag{2.41}$$

La función $D(\lambda)$ se llama **determinante característico** y la ecuación (2.41) es la **ecuación característica** de la matriz $A$. Desarrollando $D(\lambda)$ se obtiene un polinomio de $n$-ésimo grado en $\lambda$, el cual se conoce como **polinomio característico** de $A$. Por lo tanto, los eigenvalores de una matriz $A$ de $n \times n$ son las raíces del polinomio característico (de la ecuación característica, ecuación (2.41)) de grado $n$. De la teoría de polinomios se sabe que ese polinomio tiene $n$ raíces reales o complejas $\lambda_1, \lambda_2, \ldots, \lambda_n$, y algunas de las $\lambda_i$ podrían ser iguales entre sí. Por ello el polinomio se puede factorizar en la forma

$$D(\lambda) = (\lambda_1 - \lambda)^{m_1} (\lambda_2 - \lambda)^{m_2} \cdots (\lambda_r - \lambda)^{m_r} \tag{2.42}$$

siendo $\lambda_i \neq \lambda_j$ para $1 \leq i \neq j \leq r$ y $m_1 + m_2 + \cdots + m_r = n$. Al entero positivo $m_i$ se le llama la **multiplicidad algebraica** del eigenvalor $\lambda_i$. Si $m_i = 1$, al eigenvalor correspondiente se le llama un **eigenvalor simple**.

Algunas de las propiedades generales del espectro son las siguientes:

**Propiedad 1.** Dos matrices semejantes, $A$ y $\tilde{A} = S^{-1}AS$, tienen el mismo polinomio característico, lo que sigue de la igualdad

$$\begin{aligned} \det\left(\tilde{A} - \lambda I\right) &= \det\left(S^{-1}AS - \lambda I\right) = \det\left(S^{-1}AS - \lambda S^{-1} IS\right) \\ &= \det\left(S^{-1}(A - \lambda I)\,S\right) = \det(A - \lambda I) \end{aligned}$$

**Propiedad 2.** La transpuesta $A^T$ tiene los mismos eigenvalores que $A$, lo que sigue de $\det\left(A^T - \lambda I\right) = \det\left((A-\lambda I)^T\right) = \det(A - \lambda I)$.

**Propiedad 3.** Sean $\lambda_1, \lambda_2, \ldots, \lambda_n$ los eigenvalores de una matriz normal $A$ de $n \times n$, $AA^+ = A^+A$. Se demuestra que $trA = \sum_{i=1}^{n} \lambda_i$.

**Propiedad 4.** La inversa bilateral $A^{-1}$ tiene los eigenvalores $\tilde{\lambda}_1 = \lambda_1^{-1}, \tilde{\lambda}_2 = \lambda_2^{-1}, \ldots, \tilde{\lambda}_n = \lambda_n^{-1}$. Además, los eigenvectores de $A^{-1}$ son los mismos que los de $A$. Para demostrar la propiedad, considérese la ecuación característica de la matriz $A^{-1}$:

$$\begin{aligned} 0 &= \det\left(A^{-1} - \tilde{\lambda} I\right) \\ &= \det\left(\tilde{\lambda} A^{-1}\left(\tilde{\lambda}^{-1} I - A\right)\right) = \det\left(\tilde{\lambda} A^{-1}\right) \det\left(\tilde{\lambda}^{-1} I - A\right) \end{aligned}$$

Dado que $\det\left(\tilde{\lambda} A^{-1}\right) \neq 0$, se tiene que $\det\left(A^{-1} - \tilde{\lambda} I\right) = 0$ si y sólo si $\det\left(\tilde{\lambda}^{-1} I - A\right) = 0$, es decir, $\tilde{\lambda}$ es un eigenvalor de $A^{-1}$ si y sólo si $\tilde{\lambda}^{-1}$ es un eigenvalor de $A$. Además, $\mathbf{x}$ es un eigenvector de $A^{-1}$ asociado a $\tilde{\lambda}$ si y sólo si es el eigenvector de $A$ asociado a $\lambda = \tilde{\lambda}^{-1}$: $A^{1}\mathbf{x} = \tilde{\lambda}\mathbf{x} \Longleftrightarrow \tilde{\lambda}^{-1}\mathbf{x} = A\mathbf{x}$.

**Propiedad 5.** La matriz $cA$, en donde $c$ es un número, tiene un eigenvalor $\tilde{\lambda} = c\lambda$ si y sólo si $A$ tiene un eigenvalor $\lambda$. La demostración es la consecuencia inmediata de la igualdad: $0 = \det\left(cA - \tilde{\lambda} I\right) = c^n \det\left(A - \left(\tilde{\lambda}/c\right) I\right)$. Es obvio que $cA$ y $A$ tienen los mismos eigenvectores.

**Propiedad 6.** La matriz $A^m$, en donde $m$ es un entero no negativo, tiene los eigenvalores $\tilde{\lambda}_i = \lambda_i^m$, y los eigenvectores de $A$ asociados con un eigenvalor $\lambda_i$ son eigenvectores de $A^m$ asociados a $\tilde{\lambda}_i = \lambda_i^m$. La demostración es como sigue. Sea $\mathbf{x}_i$ un eigenvector de $A$ asociado a $\lambda_i$, $A\mathbf{x}_i = \lambda_i \mathbf{x}_i$. Entonces, $A^m \mathbf{x}_i = A^{m-1}(\lambda_i \mathbf{x}_i) = \ldots = \lambda_i^m \mathbf{x}_i$.

**Propiedad 7.** Una vez que se han determinado los eigenvalores, es posible determinar los eigenvectores correspondientes a partir de la ecuación (2.40). Dado que el sistema es homogéneo, es obvio que, si $\mathbf{x}$ es un eigenvector de $A$, entonces $c\mathbf{x}$, en donde $c \neq 0$ es una constante cualquiera, también es un eigenvector de $A$ correspondiente al mismo eigenvalor. En otras palabras, un eigenvector se determina hasta un factor constante.

**Ejemplo 2.14.** *La ecuación característica de la matriz*

$$A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

*en donde* $a$ *y* $b$ *son reales y* $b \neq 0$, *es* $(a - \lambda)^2 + b^2 = 0$. *Las raíces de la ecuación son complejos conjugados,* $\lambda_1 = a + ib$ *y* $\lambda_2 = a - ib$. *Los eigenvectores asociados se obtienen de sistemas de ecuaciones*

$$\begin{aligned} -ibx_1 + bx_2 &= 0 \qquad & ibx_1 + bx_2 &= 0 \\ -bx_1 - ibx_2 &= 0 \qquad & -bx_1 + ibx_2 &= 0 \end{aligned}$$

*y son*

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ i \end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ -i \end{bmatrix}$$

La situación general con respecto a la estructura del conjunto de eigenvectores es más sutil que en el caso de eigenvalores. Suponiendo conocimientos básicos del lector sobre espacios vectoriales lineales, presentamos algunas afirmaciones sobre la estructura del conjunto de todos los eigenvectores asociados.

**Teorema 2.10.** *(Eigenvectores) Sea $A$ una matriz de $n \times n$. Entonces:*

*a) existe cuando menos un eigenvector $\mathbf{x}$ asociado con cada eigenvalor distinto $\lambda$; si $A$ y $\lambda$ son reales, el eigenvector correspondiente se puede escoger real, también.*

*b) El conjunto de todos los eigenvectores $R$ asociados con un eigenvalor dado $\lambda$ junto con el vector $\mathbf{0}$ forman un subespacio invariante (propio) de $A$.*

*c) Si $\lambda_1, \lambda_2, ..., \lambda_r$ son eigenvalores diferentes, $\lambda_i \neq \lambda_j$ para $1 \leq i \neq j \leq r$, entonces, los eigenvectores $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_r$ correspondientes a diferentes eigenvalores son linealmente independientes.*

*d) Si $A$ tiene $n$ eigenvalores diferentes, entonces cualquier conjunto de $n$ eigenvectores correspondientes a esos eigenvalores es linealmente independiente, y cada eigenvector es un múltiplo de uno de esos $n$ eigenvectores.*

**Demostración.** *a) Por la definición de eigenvalor y el teorema de Cramer, existe cuando menos un eigenvector correspondiente a este eigenvalor. Si $A$ y el eigenvalor $\lambda$ son reales, y $\mathbf{x} = \mathbf{u}+i\mathbf{v}$ es un eigenvector, $A\mathbf{x} = \lambda\mathbf{x}$, pues, $A(\mathbf{u}+i\mathbf{v}) = \lambda(\mathbf{u}+i\mathbf{v})$, lo que implica que $A\mathbf{u} = \lambda\mathbf{u}$ y $A\mathbf{v} = \lambda\mathbf{v}$. Cuando menos uno de los vectores reales $\mathbf{u}$ y $\mathbf{v}$ debe ser distinto de cero, y éste puede ser escogido como el eigenvector real requerido.*

*b) El conjunto $R$ es no vacío. Sean $\mathbf{x}$ y $\mathbf{y}$ dos eigenvectores cualesquiera correspondientes a un eigenvalor $\lambda$ y $\mathbf{x},\mathbf{y} \in R$ . Entonces,*

$$A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y} = \lambda\mathbf{x} + \boldsymbol{\lambda}\mathbf{y} = \lambda(\mathbf{x}+\mathbf{y}), \quad y \quad A(\alpha\mathbf{x}) = \alpha A\mathbf{x} = \lambda(\alpha\mathbf{x})$$

*para todos los escalares $\alpha$, pues, tanto $(\mathbf{x}+\mathbf{y})$ como $\alpha\mathbf{x}$ pertenecen al conjunto $R$; es decir, para todos los elementos de $R$ se cumplen las dos operaciones básicas, adición de elementos y la multiplicación por un escalar con los postulados correspondientes, que definen un subespacio vectorial lineal. Dado que, para cualquier vector $\mathbf{x} \in R$, el vector $A\mathbf{x} \in R$, el subespacio $R$ es invariante con respecto a la matriz (transformación lineal) $A$.*

*c) Suponemos que el conjunto de eigenvectores $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_r$ es linealmente dependiente, es decir,*

$$\sum_{i=1}^{r} c_i \mathbf{x}_i = \mathbf{0} \tag{2.43}$$

*Multiplicando la matriz $(A - \lambda_1 I)$ por la igualdad anterior término a término, se tiene*

$$\sum_{i=2}^{r} c_i (\lambda_i - \lambda_1) \mathbf{x}_i = \mathbf{0}$$

*Luego, $(A - \lambda_2 I)$ se multiplica por el resultado anterior, siguiendo consecutivamente la multiplicación de $(A - \lambda_3 I)$ por el resultado hasta que llegamos a multiplicar la matriz $(A - \lambda_{r-1} I)$. El resultado de este procedimiento es la igualdad*

$$c_r (\lambda_r - \lambda_1)(\lambda_r - \lambda_2) \cdots (\lambda_r - \lambda_{r-1}) \mathbf{x}_r = \mathbf{0}$$

*de donde se tiene $c_r = 0$. Nótese que la igualdad anterior se obtiene al multiplicar la matriz*

$$C \equiv (A - \lambda_{r-1} I) \cdots (A - \lambda_2 I)(A - \lambda_1 I)$$

*por la ecuación (2.43). Dado que el orden de los términos en la combinación lineal, ecuación (2.43), no tiene importancia y cualquier término puede ser puesto al último lugar (ser el término r-ésimo), entonces se tiene que $c_1 = c_2 = \cdots = c_r = 0$, es decir, los eigenvectores $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_r$ son linealmente independientes.*

*d) Los eigenvectores $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ son linealmente independientes según el inciso c, y, por lo tanto, forman una base de espacio vectorial $C^n$. Según el inciso b, el conjunto de todos los eigenvectores $R_i$ asociados con un eigenvalor dado $\lambda_i$ junto con el vector $\mathbf{0}$ forman un subespacio invariante de $A$, y este subespacio es unidimensional. Entonces, los vectores del subespacio $R_i$ se distinguen entre sí solamente por un factor escalar, es decir, son múltiplos.* ■

## 2.7. Eigenvalores de matrices (anti)hermitianas y unitarias

Quizá reflejando las simetrías inherentes a la naturaleza, los modelos matemáticos de fenómenos del mundo real a menudo utilizan matrices hermitianas, unitarias o normales. Los eigenvalores de dichas matrices poseen las propiedades especiales resumidas en el siguiente teorema:

**Teorema 2.11.** *a) Los eigenvalores de una matriz hermitiana son reales.*

*b) Los eigenvalores de una matriz antihermitiana son imaginarios puros.*

*c) Los eigenvalores de una matriz unitaria tienen el valor absoluto igual a uno.*

**Demostración.** *a) Sea $\lambda$ un eigenvalor de una matriz hermitiana $A$, $A^+ = A$. Por definición, existe un eigenvector $\mathbf{x} \neq \mathbf{0}$ tal que $A\mathbf{x} = \lambda\mathbf{x}$. La conjugada hermitiana de la igualdad anterior es $\mathbf{x}^+ A = \lambda^* \mathbf{x}^+$. Multiplicando la primera igualdad por $\mathbf{x}^+$ por la izquierda y la segunda igualdad por $\mathbf{x}$ por la derecha, se obtienen $\mathbf{x}^+ A\mathbf{x} = \lambda \mathbf{x}^+\mathbf{x}$ y $\mathbf{x}^+ A\mathbf{x} = \lambda^* \mathbf{x}^+\mathbf{x}$. Restando las dos igualdades y tomando en cuenta que el producto escalar $\mathbf{x}^+\mathbf{x} > 0$, se tiene que $\lambda - \lambda^* = 0$, es decir, el eigenvalor $\lambda$ es real.*

*b) De manera similar, para una matriz antihermitiana $A$, $A^+ = -A$, que tiene un eigenvalor $\lambda$, se obtienen dos igualdades $\mathbf{x}^+ A\mathbf{x} = \lambda \mathbf{x}^+\mathbf{x}$ y $-\mathbf{x}^+ A\mathbf{x} =$*

$\lambda^* \mathbf{x}^+ \mathbf{x}$. *Sumándolas, se tiene que* $\lambda + \lambda^* = 0$, *es decir, el eigenvalor* $\lambda$ *es imaginario puro.*

*c) Sea* $\lambda$ *un eigenvalor y* $\mathbf{x} \neq \mathbf{0}$ *un eigenvector correspondiente de una matriz unitaria* $U$, $U^+ = U^{-1}$. *Entonces,* $U\mathbf{x} = \lambda \mathbf{x}$ *y* $\mathbf{x}^+ U^{-1} = \lambda^* \mathbf{x}^+$. *Multiplicando por partes las igualdades, se tiene* $\mathbf{x}^+ U^{-1} U \mathbf{x} = (\lambda^* \mathbf{x}^+)(\lambda \mathbf{x})$, *esto es,* $\mathbf{x}^+ \mathbf{x} = (\lambda^* \lambda)\, \mathbf{x}^+ \mathbf{x}$. *De donde,* $\lambda^* \lambda = |\lambda|^2 = 1$ *y queda completa la demostración.* ■

Resulta que para el análisis del sistema de eigenvectores de matrices especiales consideradas anteriormente es útil el concepto de matrices normales.

**Definición 2.19.** *Una matriz* $A$ *de* $n \times n$ *recibe el nombre de* ***matriz normal*** *si es conmutativa con su conjugada hermitiana,* $AA^+ = A^+ A$.

Es fácil ver que las matrices (anti)hermitianas y unitarias pertenecen a la clase de matrices normales, es decir, son casos particulares de matrices normales.

**Teorema 2.12.** *Una matriz cuadrada* $A$ *puede ser siempre representada en la forma*

$$A = H_1 + iH_2 \tag{2.44}$$

*en donde* $H_1$ *y* $H_2$ *son dos matrices hermitianas unívocamente determinadas por* $A$,

$$H_1 = \frac{1}{2}\left(A + A^+\right), \quad H_2 = \frac{1}{2i}\left(A - A^+\right) \tag{2.45}$$

*Además, la matriz* $A$ *es normal si y sólo si sus componentes hermitianos* $H_1$ *y* $H_2$ *son conmutativos.*

**Demostración.** *Es obvio que las dos matrices* $H_1$ *y* $H_2$, *ecuación (2.45), son hermitianas y unívocamente determinadas por* $A$ *y, además, es válida la representación (2.44).*

*Por otra parte, si* $A$ *es una matriz normal,* $AA^+ = A^+ A$, *entonces, usando la ecuación (2.44) se tiene* $H_1 H_2 = H_2 H_1$. *Y viceversa, si* $H_1 H_2 = H_2 H_1$, *entonces, de la ecuación (2.44) se sigue que* $AA^+ = A^+ A$, *es decir,* $A$ *es normal. Quedando completa la demostración.* ■

Los siguientes teoremas se enuncian sin demostración. Las demostraciones véanse, por ejemplo, en [8] o [16].

**Teorema 2.13.** *Una matriz* $A$ *de* $n \times n$ *es normal si y sólo si tiene un conjunto de* $n$ *eigenvectores ortonormales.*

Sea $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ un conjunto ortonormal de $n$ eigenvectores (columnas) de una matriz normal $A$ correspondientes a los eigenvalores $\lambda_1, \lambda_2, \ldots, \lambda_n$. La matriz $U = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]$ es unitaria y, además, $AU = [\lambda_1 \mathbf{x}_1, \lambda_2 \mathbf{x}_2, \ldots, \lambda_n \mathbf{x}_n]$, $U^+ AU = U^+ [\lambda_1 \mathbf{x}_1, \lambda_2 \mathbf{x}_2, \ldots, \lambda_n \mathbf{x}_n] = diag\,[\lambda_1, \lambda_2, \ldots, \lambda_n]$, en donde $diag\,[\lambda_1, \lambda_2, \ldots, \lambda_n]$ es la matriz diagonal de eigenvalores. Por tanto, el teorema anterior es equivalente al siguiente.

**Teorema 2.14.** *Una matriz $A$ de $n \times n$ es normal si y sólo si es unitariamente semejante a una matriz diagonal,*

$$U^{+}AU = diag\,[\lambda_1, \lambda_2, \ldots, \lambda_n]\,, \quad U^{+} = U^{-1} \tag{2.46}$$

*Los elementos en la diagonal principal son, por supuesto, eigenvalores de $A$.*

Del teorema anterior y del teorema de eigenvalores de una matriz hermitiana se sigue:

**Teorema 2.15.** *Una matriz $A$ de $n \times n$ es hermitiana si y sólo si es unitariamente semejante a una matriz diagonal con elementos reales,*

$$U^{+}AU = diag\,[\lambda_1, \lambda_2, \ldots, \lambda_n]\,, \quad U^{+} = U^{-1}, \quad \lambda_i = \lambda_i^{*} \quad (i = 1, 2, \ldots, n) \tag{2.47}$$

**Teorema 2.16.** *Una matriz $A$ de $n \times n$ es unitaria si y sólo si es unitariamente semejante a una matriz diagonal con elementos que tienen el valor absoluto igual a uno en la diagonal principal,*

$$U^{+}AU = diag\,[\lambda_1, \lambda_2, \ldots, \lambda_n]\,, \quad U^{+} = U^{-1}, \quad |\lambda_i| = 1 \quad (i = 1, 2, \ldots, n) \tag{2.48}$$

**Problema.** *Sea $A$ una matriz cuadrada. Demuestre que $\lambda$ es un eigenvalor de $A$ si y sólo si $\lambda^{*}$ es un eigenvalor de $A^{+}$. Sugerencia: considérese la ecuación característica* $\det(A - \lambda I) = (\det(A^{+} - \lambda^{*} I))^{*} = 0$.

**Problema.** *Sea $A$ una matriz normal, $AA^{+} = A^{+}A$. Demuestre que si $\mathbf{x}$ es un eigenvector de $A$ asociado con un eigenvalor $\lambda$, $A\mathbf{x} = \lambda\mathbf{x}$, y a la vez $\mathbf{x}$ es un eigenvector de $A^{+}$ asociado con un eigenvalor $\mu$, $A^{+}\mathbf{x} = \mu\mathbf{x}$, entonces, $\mu = \lambda^{*}$.*

**Problema.** *Sean $\lambda_1, \lambda_2, \ldots, \lambda_n$ los eigenvalores de una matriz normal $A$ de $n \times n$, $AA^{+} = A^{+}A$. Demuestre que $tr A = \sum_{i=1}^{n} \lambda_i$.*

## 2.8. Función de una matriz

Sean $A = [a_{ij}]$ una matriz cuadrada de $n \times n$ y $f(\lambda)$ una función de variable escalar $\lambda$. Es natural el deseo de extender el concepto de una función escalar a una función de variable matricial. Dado que anteriormente fue definida la potencia entera positiva $A^n$ de una matriz, la solución de dicho problema es fácil en el caso de funciones que son polinomios de su variable $\lambda$; es decir, $f(\lambda) = a_0\lambda^m + a_1\lambda^{m-1} + \cdots + a_m$ tiene como su homólogo la siguiente función de matriz $A$, $f(A) = a_0A^m + a_1A^{m-1} + \cdots + a_mA^0$. Sin embargo, el problema general de función de una matriz es algo complicado y requiere conocimientos de la teoría espectral de matrices. Por esta razón, en esta sección las consideraciones serán un tanto reducidas. Una manera práctica de introducir el concepto de función de una matriz es como sigue.

**Definición 2.20.** *Sean $A = [a_{ij}]$ una matriz cuadrada de $n \times n$ y $f(\lambda)$ una función de variable escalar $\lambda$ con su desarrollo de Taylor*

$$f(\lambda) = \sum_{m=0}^{\infty} \frac{f^{(m)}(0)}{m!} \lambda^m$$

*Entonces, la función matricial se define como la serie de potencias de la matriz $A$,*

$$f(A) = \sum_{m=0}^{\infty} \frac{f^{(m)}(0)}{m!} A^m \tag{2.49}$$

*en donde $A^0 = I$. Se supone que la serie de Taylor de la función $f(\lambda)$ converge para cada uno de los eigenvalores de la matriz $A$.*

Nótese que la función $f(A)$ de una matriz $A$ de $n \times n$ es una matriz cuadrada del orden de $A$.

**Ejemplo 2.15.** *La definición anterior, por supuesto, contiene como un caso particular funciones polinomiales. Otras funciones matriciales de interés práctica son:*

$$\begin{aligned}
\exp A &= \sum_{m=0}^{\infty} \frac{A^m}{m!}, \quad \cos A = \frac{e^{iA} + e^{-iA}}{2} = \sum_{m=0}^{\infty} \frac{(-1)^m}{(2m)!} A^{2m} \\
\operatorname{sen} A &= \frac{e^{iA} - e^{-iA}}{2i} = \sum_{m=0}^{\infty} \frac{(-1)^m}{(2m+1)!} A^{2m+1} \\
\cosh A &= \frac{e^{A} + e^{-A}}{2} = \sum_{m=0}^{\infty} \frac{A^{2m}}{(2m)!}, \quad \operatorname{senh} A = \frac{e^{A} - e^{-A}}{2} = \sum_{m=0}^{\infty} \frac{A^{2m+1}}{(2m+1)!} \\
(I - A)^{-1} &= \sum_{m=0}^{\infty} A^m, \quad \textit{cuando } |\lambda_k| < 1; \quad k = 1, 2, \ldots, s \\
\ln A &= \sum_{m=1}^{\infty} \frac{(-1)^{m-1}}{m} (A - I)^m, \quad \textit{cuando } |\lambda_k - 1| < 1; \quad k = 1, 2, \ldots, s
\end{aligned}$$

*donde $\lambda_k$ son eigenvalores de $A$.*

Algunas de las propiedades de funciones de una matriz son las siguientes.

**Propiedad 1.** Sea $\mathbf{x}$ un eigenvector de $A$ correspondiente al eigenvalor $\lambda$, $A\mathbf{x} = \lambda\mathbf{x}$. Entonces, $f(A)\mathbf{x} = f(\lambda)\mathbf{x}$, es decir, $\mathbf{x}$ es eigenvector de la función matricial $f(A)$ correspondiente al eigenvalor $f(\lambda)$. (Compruébelo)

**Propiedad 2.** Si $A$ y $B$ son dos matrices semejantes y $S$ es la matriz de transformación, $B = S^{-1}AS$, entonces, las matrices $f(A)$ y $f(B)$ son también semejantes con la misma matriz de transformación, $f(B) = S^{-1}f(A)S$. (Compruébelo)

Las siguientes igualdades se presentan sin demostración.

**Propiedad 3.** Si $A$ y $B$ son dos matrices conmutativas, $AB = BA$, entonces,

$$e^A e^B = e^B e^A = e^{A+B} \tag{2.50}$$

de donde, para cualquier matriz $A$, se tiene

$$e^O = I, \quad e^A e^{-A} = I, \quad e^{-A} = \left(e^A\right)^{-1} \tag{2.51}$$

$$e^{iA} = \cos A + i \operatorname{sen} A, \quad \cos^2 A + \operatorname{sen}^2 A = I \tag{2.52}$$

## 2.9. Sistemas de ecuaciones diferenciales lineales

La solución de un sistema de ecuaciones diferenciales lineales con coeficientes constantes se puede presentar en forma matricial. Con este fin, primero definimos la derivada de una matriz de funciones.

**Definición 2.21.** *Si los elementos $a_{ij}(t)$ de una matriz $A(t) = [a_{ij}(t)]$ son funciones de una variable $t$, se define la derivada de ésta como la matriz de derivadas:*

$$\frac{dA(t)}{dt} = \left[\frac{da_{ij}(t)}{dt}\right] \tag{2.53}$$

Basándose en la definición, es fácil ver que si la matriz $AB$ está definida, entonces

$$\frac{d(AB)}{dt} = A\frac{dB}{dt} + \frac{dA}{dt}B \tag{2.54}$$

Además, si $A$ es no singular, entonces, diferenciando $AA^{-1} = I$, se tiene

$$\frac{d\left(A^{-1}\right)}{dt} = -A^{-1}\frac{dA}{dt}A^{-1} \tag{2.55}$$

Un sistema homogéneo de ecuaciones diferenciales lineales de primer orden con coeficientes constantes se puede representar en forma matricial como sigue:

$$\frac{d\mathbf{x}(t)}{dt} = A\mathbf{x}(t) \tag{2.56}$$

en donde $t$ es la variable independiente, $\mathbf{x}(t)$ es un vector columna de $n$ incógnitas que son funciones de $t$, $\mathbf{x}^T(t) = [x_1(t), x_2(t), \ldots, x_n(t),]$, $A = [a_{ij}]$ es una matriz cuadrada de $n \times n$ de constantes complejas. Puesto que las condiciones iniciales son

$$\mathbf{x}(0) = \mathbf{x}_0, \quad \text{es decir,} \quad x_1(0) = x_{10},\ x_2(0) = x_{20}, \ldots, x_n(0) = x_{n0} \tag{2.57}$$

Buscaremos la solución del sistema en forma de serie de Taylor en potencias de la variable $t$:

$$\mathbf{x}(t) = \mathbf{x}_0 + \dot{\mathbf{x}}_0 t + \ddot{\mathbf{x}}_0\frac{t^2}{2!} + \cdots, \text{ en donde } \dot{\mathbf{x}}_0 = \frac{dx(0)}{dt}, \quad \ddot{\mathbf{x}}_0 = \frac{d^2x(0)}{dt^2}, \ \ldots \tag{2.58}$$

Diferenciando la ecuación (2.56), se tiene que

$$\frac{d^2\mathbf{x}(t)}{dt^2} = A\dot{\mathbf{x}}(t) = A^2\mathbf{x}(t), \quad \frac{d^3\mathbf{x}(t)}{dt^3} = A\ddot{\mathbf{x}}(t) = A^3\mathbf{x}(t), \ \ldots$$

Luego, de la última ecuación y la ecuación (2.56) se obtienen las derivadas de $\mathbf{x}(t)$ al instante $t = 0$:

$$\dot{\mathbf{x}}_0 = A\mathbf{x}_0, \quad \ddot{\mathbf{x}}_0 = A^2\mathbf{x}_0, \ \ldots$$

Sustituyéndolas en la ecuación (2.58) se tiene que

$$\mathbf{x}(t) = \mathbf{x}_0 + tA\mathbf{x}_0 + \frac{t^2}{2!}A^2\mathbf{x}_0 + \cdots = e^{tA}\mathbf{x}_0 \tag{2.59}$$

es la solución buscada del sistema homogéneo de ecuaciones diferenciales de primer orden con coeficientes constantes.

**Problema.** *Demuestre que la ecuación (2.59) es la solución de la ecuación (2.56) mediante la sustitución directa de la (2.59) en la ecuación (2.56).*

**Problema.** *Suponiendo que las condiciones iniciales del sistema están dadas en un instante $t_0$, demuestre que la solución es $\mathbf{x}(t) = e^{(t-t_0)A}\mathbf{x}_0$.*

**Problema.** *Sea $A(t)$ una matriz cuadrada de $n \times n$ de una variable independiente $t$ y $f(\lambda)$ una función de variable escalar $\lambda$. Por tanto, la matriz $B(t) \equiv f(A(t))$ es una matriz de $n \times n$ de la variable $t$. ¿Bajo qué condiciones es válida la regla de la cadena para $B(t)$: $dB(t)/dt = (df(A)/dA)(dA(t)/dt)$?, en donde se entiende que $df(A)/dA = (df(\lambda)/d\lambda)\,|_{\lambda=A}$. (Sugerencia: considérese la función $f(\lambda) = \lambda^2$.)*

En el caso de un sistema no homogéneo de ecuaciones diferenciales de primer orden con coeficientes constantes,

$$\frac{d\mathbf{x}(t)}{dt} = A\mathbf{x}(t) + \mathbf{f}(t) \tag{2.60}$$

en donde $\mathbf{f}(t)$ es un vector columna de funciones conocidas $f_1(t), f_2(t), \ldots, f_n(t)$, la solución se busca en forma de

$$\mathbf{x}(t) = e^{tA}\mathbf{z}(t) \tag{2.61}$$

Sustituyendo en la ecuación (2.60), se tiene la ecuación

$$e^{tA}\frac{d\mathbf{z}(t)}{dt} = \mathbf{f}(t)$$

que tiene como su solución el vector columna

$$\mathbf{z}(t) = \mathbf{c} + \int_{t_0}^{t} e^{-\tau A}\mathbf{f}(\tau)\,d\tau$$

en donde $\mathbf{c}$ es un vector columna de constantes de integración y la integral de una matriz se entiende como la matriz de integrales de los elementos. Luego,

$$\mathbf{x}(t) = e^{tA}\mathbf{z}(t) = e^{tA}\left(\mathbf{c} + \int_{t_0}^{t} e^{-\tau A}\mathbf{f}(\tau)\, d\tau\right) = e^{tA}\mathbf{c} + \int_{t_0}^{t} e^{(t-\tau)A}\mathbf{f}(\tau)\, d\tau$$

Las condiciones iniciales para $\mathbf{x}(t)$ al instante inicial $t_0$ requieren que se cumpla la igualdad

$$\mathbf{x}_0 = \mathbf{x}(t_0) = e^{t_0 A}\mathbf{c}, \quad \text{es decir,} \quad \mathbf{c} = e^{-t_0 A}\mathbf{x}_0$$

Finalmente, la solución vectorial de la ecuación (2.60) es

$$\mathbf{x}(t) = e^{(t-t_0)A}\mathbf{x}_0 + \int_{t_0}^{t} e^{(t-\tau)A}\mathbf{f}(\tau)\, d\tau \tag{2.62}$$

## 2.10. Formas bilineales, cuadráticas y hermitianas

Las funciones cuadráticas de múltiples variables independientes son del siguiente nivel de complejidad después de las constantes y funciones lineales. Tales funciones surgen en distintas áreas de aplicación y los métodos matriciales permiten un estudio unificado de sus propiedades. Por otra parte, las funciones cuadráticas proporcionan conceptos matriciales importantes, especialmente en lo que se refiere a los sistemas de eigenvalores y eigenvectores. Las curvas llamadas secciones cónicas, que juegan un papel fundamental en geometría bidimensional, son un ejemplo conocido de las funciones cuadráticas. Muchos problemas de optimización y estabilidad de sistemas dinámicos llevan a la necesidad de estudiar el comportamiento de un sistema en vecindades de estados de equilibrio. Tales estudios se basan en una función de múltiples variables que puede llamarse función de costo o energía. Los valores extremos y el comportamiento de esta función en vecindades de los estados de equilibrio a menudo se aproximan por medio de funciones cuadráticas, desarrollando la función de costo en serie de Taylor. Cabe mencionar que las formas bilineales y cuadráticas hermitianas son de importancia especial para la construcción de espacios unitarios en Mecánica Cuántica.

Un polinomio homogéneo de segundo orden con respecto a las $2n$ variables $x_i$ y $y_j$ $(i, j = 1, 2, \ldots, n)$,

$$A(\mathbf{x}, \mathbf{y}) = \sum_{i,j=1}^{n} a_{ij} x_i^* y_j \tag{2.63}$$

se llama **forma bilineal** de las $2n$ variables $x_i$ y $y_j$. Las $2n$ variables $x_i$ y $y_j$

pueden considerarse como componentes de dos vectores columna

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

a la vez que los números $a_{ij}$ pueden considerarse como elementos de una matriz $A = [a_{ij}]$ de $n \times n$. Por lo tanto, la forma bilineal puede escribirse en forma matricial

$$A(\mathbf{x}, \mathbf{y}) = \mathbf{x}^+ A \mathbf{y} \tag{2.64}$$

En general, tanto las variables $x_i$ y $y_j$ como los números $a_{ij}$ pueden ser complejos.

**Ejemplo 2.16.** *Si $A$ es la matriz de unidad, $A = I$, y $\mathbf{x}$ y $\mathbf{y}$ son vectores reales, entonces*

$$A(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T I \mathbf{y} = \mathbf{x}^T \mathbf{y} = (\mathbf{x}, \mathbf{y})$$

*es el producto interior de los vectores $\mathbf{x}$ y $\mathbf{y}$.*

A cada forma bilineal le corresponde una **forma cuadrática** tomando $\mathbf{y} = \mathbf{x}$ en la ecuación (2.64),

$$A(\mathbf{x}, \mathbf{x}) = \mathbf{x}^+ A \mathbf{x} = \sum_{i,j=1}^{n} a_{ij} x_i^* x_j \tag{2.65}$$

de las $n$ variables $x_i$ posiblemente complejas. Las formas cuadráticas, que sólo toman valores reales, son generalmente las más requeridas en aplicaciones. El siguiente teorema contesta a la pregunta: ¿bajo qué condiciones las formas cuadráticas son reales?

**Teorema 2.17.** *a) La forma cuadrática $A(\mathbf{x}, \mathbf{x})$ generada por una matriz hermitiana $A$ de $n \times n$ es real para toda $\mathbf{x}$ compleja o real.*

*b) $A(\mathbf{x}, \mathbf{x})$ es real para toda $\mathbf{x}$ real si $A = R + H$, en donde $R$ es real y $H$ es hermitiana.*

*c) Para toda $\mathbf{x}$ real la forma cuadrática $A(\mathbf{x}, \mathbf{x})$ generada por cualquier matriz $A$ es idéntica a la generada por la matriz simétrica $C = \left(A + A^T\right)/2$.*

**Demostración.** *a) Si $A$ es hermitiana (o real simétrica), entonces*

$$\begin{aligned} \left(\mathbf{x}^+ A \mathbf{x}\right)^* &= \left(\sum_{i,j=1}^{n} a_{ij} x_i^* x_j\right)^* = \sum_{i,j=1}^{n} a_{ij}^* x_j^* x_i \\ &= \sum_{i,j=1}^{n} a_{ji}^+ x_j^* x_i = \mathbf{x}^+ A^+ \mathbf{x} = \mathbf{x}^+ A \mathbf{x} \end{aligned}$$

*Por lo tanto, la forma cuadrática* $A(\mathbf{x},\mathbf{x})$ *es real para cualquier matriz hermitiana.*

*b) Sea* $A = R + H$, *en donde* $R$ *es real y* $H$ *es hermitiana. La forma* $H(\mathbf{x},\mathbf{x}) = \mathbf{x}^{+}H\mathbf{x} = \mathbf{x}^{T}H\mathbf{x}$ *es real para toda* $\mathbf{x}$. *Si* $R$ *es una matriz real, entonces para toda* $\mathbf{x}$ *real la forma cuadrática* $R(\mathbf{x},\mathbf{x})$ *es real. Por lo tanto,* $A(\mathbf{x},\mathbf{x}) = \mathbf{x}^{T}(R+H)\mathbf{x} = \mathbf{x}^{T}R\mathbf{x} + \mathbf{x}^{T}H\mathbf{x}$ *es también real.*

*c) Nótese que para toda* $\mathbf{x}$ *real y cualquier matriz* $A$, *las formas cuadráticas* $A(\mathbf{x},\mathbf{x})$ *y* $A^{T}(\mathbf{x},\mathbf{x})$, *generadas por* $A$ *y* $A^{T}$, *respectivamente, son idénticas:*

$$
\begin{aligned}
A(\mathbf{x},\mathbf{x}) &= \mathbf{x}^{T}A\mathbf{x} = \sum_{i,j=1}^{n} a_{ij}x_i x_j = \sum_{i,j=1}^{n} a_{ji}x_j x_i \\
&= \sum_{i,j=1}^{n} a_{ij}^{T}x_i x_j = \mathbf{x}^{T}A^{T}\mathbf{x} = A^{T}(\mathbf{x},\mathbf{x})
\end{aligned}
$$

*De donde, representando* $A = \left(A + A^{T}\right)/2 + \left(A - A^{T}\right)/2$, *se tiene que*

$$
\begin{aligned}
A(\mathbf{x},\mathbf{x}) &= \mathbf{x}^{T}A\mathbf{x} = \mathbf{x}^{T}\left(\left(A + A^{T}\right)/2 + \left(A - A^{T}\right)/2\right)\mathbf{x} \\
&= \mathbf{x}^{T}\left(\left(A + A^{T}\right)/2\right)\mathbf{x}
\end{aligned}
$$

*Esto completa la demostración.* ■

**Teorema 2.18.** *Para toda* $\mathbf{x}$ *compleja o real, la forma cuadrática* $A(\mathbf{x},\mathbf{x})$ *generada por una matriz antihermitiana* $A$, $A^{+} = -A$, *es un número imaginario puro o cero.*

**Demostración.** *Si* $A$ *es antihermitiana, entonces*

$$
\begin{aligned}
\left(\mathbf{x}^{+}A\mathbf{x}\right)^{*} &= \left(\sum_{i,j=1}^{n} a_{ij}x_i^{*}x_j\right)^{*} = \sum_{i,j=1}^{n} a_{ij}^{*}x_j^{*}x_i \\
&= \sum_{i,j=1}^{n} a_{ji}^{+}x_j^{*}x_i = \mathbf{x}^{+}A^{+}\mathbf{x} = -\mathbf{x}^{+}A\mathbf{x}
\end{aligned}
$$

*así, el valor de la forma cuadrática es un número imaginario puro o cero.* ■

**Nota bibliográfica:** para ampliar los conocimientos sobre el tema de este capítulo, puede consultar [1], [3], [7], [8], [9], [10], [11] y [16].

# Capítulo 3

# Variable compleja

Muchos problemas de física e ingeniería pueden tratarse por medio de métodos de análisis complejo. En términos matemáticos, estos problemas pueden dividirse en dos clases: la primera abarca problemas elementales que requieren conocimiento acerca de números complejos y funciones complejas elementales; la segunda requiere conocimientos a detalle de la teoría de las funciones analíticas. Los ejemplos de problemas de la primera clase están relacionados con circuitos eléctricos y sistemas vibrantes. Varios problemas interesantes de la dinámica de fluidos, la teoría del calor y la electrostática pertenecen a la segunda categoría.

La importancia de las funciones analíticas está dictada principalmente por tres razones. Primera, muchas integrales complicadas, reales y complejas, que aparecen en aplicaciones se pueden evaluar por los métodos de la integración compleja. Segunda, las partes real e imaginaria de una función analítica son soluciones de la ecuación de Laplace en dos variables independientes. Por tanto, los problemas bidimensionales de potencial pueden tratarse mediante los métodos de las funciones analíticas. Tercera, las transformaciones o mapeos conformes de regiones de planos complejos representan una herramienta poderosa de simplificación de la solución de problemas definidos inicialmente en dominios complicados. Además, la mayoría de las funciones no elementales de las Matemáticas son funciones analíticas y, por tanto, muchas de sus propiedades se descubren al considerar éstas para valores complejos de la variable independiente. Las secciones que siguen proporcionan las bases tanto de números complejos como de las funciones analíticas y sus aplicaciones.

## 3.1. Números complejos

### 3.1.1. Definición y álgebra de números complejos

En el estudio de las Matemáticas se puede observar que existen muchas ecuaciones algebraicas que no se resuelven por ningún número real, por ejemplo,

$$x^2 + 1 = 0, \quad x^4 + x^2 + 2 = 0$$

Esto conduce a la necesidad de la introducción de los números complejos. Si adjuntamos a los números reales un símbolo $i$ tal que satisface por definición a la ecuación

$$i^2 = -1 \tag{3.1}$$

podemos construir los números complejos,

$$z = x + iy = \operatorname{Re} z + i \operatorname{Im} z \tag{3.2}$$

en donde $x$ y $y$ son números reales cualesquiera. A $\operatorname{Re} z = x$ se le llama **parte real** del número complejo $z$, e $\operatorname{Im} z = y$ recibe el nombre de **parte imaginaria** de $z$. Resulta que los números complejos satisfacen las leyes algebraicas de los números reales e incluyen a los números reales como un caso especial. Además, los números complejos permiten resolver no solamente la ecuación $x^2 + 1 = 0$, sino cualquier ecuación polinomial.

Se dice que dos números complejos $z_1 = x_1 + iy_1$ y $z_2 = x_2 + iy_2$ son iguales si y sólo si sus partes reales e imaginarias son respectivamente iguales, es decir, $x_1 = x_2$ y $y_1 = y_2$. Además, se admite que $iy = yi$. La adición y la multiplicación de números complejos se definen de tal manera que el álgebra de números complejos sea similar a la de los números reales.

La **suma (adición) de dos números complejos** $z_1 = x_1 + iy_1$ y $z_2 = x_2 + iy_2$ es

$$z_1 + z_2 = (x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2) \tag{3.3}$$

**El producto (multiplicación) de dos números complejos** $z_1$ y $z_2$ se define como

$$\begin{aligned} z_1 \cdot z_2 &= (x_1 + iy_1)(x_2 + iy_2) = x_1x_2 + iy_1x_2 + ix_1y_2 + i^2y_1y_2 \\ &= (x_1x_2 - y_1y_2) + i\,(x_1y_2 + y_1x_2) \end{aligned} \tag{3.4}$$

A partir de las leyes del álgebra para los números reales se obtienen las leyes para los números complejos. La **ley conmutativa** de la adición y la multiplicación:

$$z_1 + z_2 = z_2 + z_1, \qquad z_1 \cdot z_2 = z_2 \cdot z_1 \tag{3.5}$$

La **ley asociativa** de la adición y la multiplicación:

$$(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3), \quad (z_1 \cdot z_2) \cdot z_3 = z_1 \cdot (z_2 \cdot z_3) \tag{3.6}$$

La **ley distributiva**:

$$z_1(z_2 + z_3) = z_1z_2 + z_1z_3 \tag{3.7}$$

Además, se define el número cero $\widehat{0}$ y la unidad $\widehat{1}$ tales que para cualquier número complejo $z$ se cumplan las igualdades,

$$z + \widehat{0} = z, \qquad z \cdot \widehat{1} = z \tag{3.8}$$

Es fácil ver que $\widehat{0} = 0 + i0 = 0$ y $\widehat{1} = 1 + i0 = 1$.

La **sustracción** se define como la inversa de la adición, es decir, la diferencia $z_1 - z_2$ es un número complejo $z$ tal que $z_1 = z + z_2$. Es obvio que

$$z_1 - z_2 = (x_1 - x_2) + i\,(y_1 - y_2) \tag{3.9}$$

La **división** se define como la inversa de la multiplicación, es decir, el cociente $z_1/z_2 = z$ es el número complejo $z$ para el que $z_1 = zz_2$. La última ecuación puede escribirse como

$$(x_1 + iy_1) = (x + iy)\,(x_2 + iy_2)$$

Igualando las partes reales y las imaginarias, se tiene un sistema de dos ecuaciones lineales en las incógnitas $x$ y $y$:

$$\begin{aligned} x_1 &= xx_2 - yy_2 \\ x_1 &= xy_2 + yx_2 \end{aligned}$$

Suponiendo $z_2 \neq 0$, se obtiene la solución única

$$x = \frac{x_1x_2 + y_1y_2}{x_2^2 + y_2^2} \quad y = \frac{y_1x_2 - x_1y_2}{x_2^2 + y_2^2}$$

Nótese que la manera práctica de obtener el cociente es como sigue:

$$\begin{aligned} z &= \frac{z_1}{z_2} = \frac{x_1 + iy_1}{x_2 + iy_2} = \frac{(x_1 + iy_1)\,(x_2 - iy_2)}{(x_2 + iy_2)\,(x_2 - iy_2)} \\ &= \frac{x_1x_2 + y_1y_2}{x_2^2 + y_2^2} + i\frac{y_1x_2 - x_1y_2}{x_2^2 + y_2^2} \end{aligned} \tag{3.10}$$

Para todo número complejo $z$ se define su **complejo conjugado**,

$$z^* = x - iy = x + i(-y) \tag{3.11}$$

Aplicando las reglas de la adición, se tiene que $z + z^* = 2\,\mathrm{Re}\,z$ y $z - z^* = 2\,\mathrm{Im}\,z$. Entonces, $z$ es un número real si y sólo si $z^* = z$ y $z$ es un número imaginario puro si y sólo si $z^* = -z$. Además, $(z_1 + z_2)^* = z_1^* + z_2^*$.

### 3.1.2. Representación vectorial y polar de un número complejo. Fórmula de De Moivre

Como se puede ver de la sección anterior, un número complejo $z = x + iy$ se determina unívocamente por el par ordenado de números reales $(x, y)$. Esto nos da la idea de una representación geométrica de los números complejos como puntos o vectores en el plano, que tiene una gran importancia en las aplicaciones. En un sistema de coordenadas cartesianas, al eje $x$ se le da el nombre de eje real y al eje $y$ se le llama el eje imaginario. Luego, un número complejo $z = x + iy$ se representa como el punto o vector con coordenadas $(x, y)$ (véase la figura 3.1). Al plano $xy$ en el que se representan los números complejos se le llama **plano complejo**. El **valor absoluto** o **módulo** de $z$ es la distancia

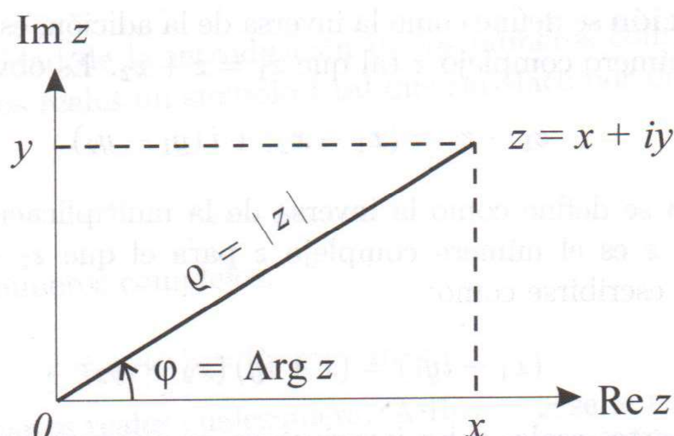



Figura 3.1. Plano complejo, presentación vectorial y polar de un número complejo

del punto de representación al origen y se denota mediante $| z |$. Es claro que $| z |= \sqrt{x^2 + y^2} = \sqrt{zz^*}$. Nótese que la adición de dos números complejos se implementa en el plano complejo como la adición de dos vectores correspondientes, según la regla del paralelogramo.

El cambio del sistema de coordenadas cartesianas al sistema polar,

$$x = \rho \cos \varphi, \quad y = \rho \operatorname{sen} \varphi \tag{3.12}$$

permite escribir un número complejo $z = x + iy$ en la forma

$$z = x + iy = \rho\,(\cos \varphi + i \operatorname{sen} \varphi) \tag{3.13}$$

que se conoce como la **forma polar** o **trigonométrica** de un número complejo. En esta representación $\rho = \sqrt{x^2 + y^2}$ es el módulo de $z$ y el ángulo $\varphi$, medido en radianes y considerado positivo al medirlo en sentido contrario al de las manecillas del reloj, se llama **argumento** de $z$ y se denota como $\arg z = \varphi = \arctan{(y/x)}$. Para $z = 0$ este ángulo no está definido. Para $z \neq 0$ está determinado sólo hasta múltiplos enteros de $2\pi$. Para evitar ambigüedades, a menudo se utiliza el **valor principal del argumento**, el cual se define por medio de $0 < \varphi < 2\pi$.

La interpretación geométrica de la multiplicación y la división de números complejos es más fácil en la forma polar. En efecto, en virtud de fórmulas trigonométricas conocidas se tiene

$$\begin{aligned} z_1 \cdot z_2 &= \rho_1\,(\cos \varphi_1 + i \operatorname{sen} \varphi_1)\, \rho_2\,(\cos \varphi_2 + i \operatorname{sen} \varphi_2) \\ &= \rho_1 \rho_2\,(\cos{(\varphi_1 + \varphi_2)} + i \operatorname{sen}{(\varphi_1 + \varphi_2)}) \end{aligned} \tag{3.14}$$

Por inducción, es fácil extender la fórmula anterior a cualquier número de multiplicadores, es decir,

$$z_1 \cdot z_2 \cdot \cdots \cdot z_n = \rho_1 \rho_2 \cdots \rho_n \Big(\cos\Big(\sum \varphi_i\Big) + i \operatorname{sen}\Big(\sum \varphi_i\Big)\Big) \tag{3.15}$$

Entonces, el módulo del producto es igual al producto de los módulos y el argumento del producto es igual a la suma de los argumentos.

Un caso particular de la ecuación (3.15) es la **fórmula de De Moivre**

$$z^n = \rho^n (\cos n\varphi + i \operatorname{sen} n\varphi) \tag{3.16}$$

en donde $n$ es un número natural.

La división de números complejos en forma polar se da mediante la fórmula

$$\frac{z_1}{z_2} = \frac{\rho_1 (\cos\varphi_1 + i \operatorname{sen}\varphi_1)}{\rho_2 (\cos\varphi_2 + i \operatorname{sen}\varphi_2)} = \frac{\rho_1}{\rho_2} (\cos(\varphi_1 - \varphi_2) + i \operatorname{sen}(\varphi_1 - \varphi_2)) \tag{3.17}$$

de la cual se sigue

$$z^{-1} = \frac{1}{z} = \frac{\cos 0 + i \operatorname{sen} 0}{\rho (\cos\varphi + i \operatorname{sen}\varphi)} = \frac{1}{\rho} (\cos(-\varphi) + i \operatorname{sen}(-\varphi))$$

De la última igualdad y la fórmula de De Moivre se establece la fórmula para las potencias enteras negativas de un número complejo $z$:

$$z^{-n} = \left(\frac{1}{z}\right)^n = \left(\frac{1}{\rho}\right)^n (\cos(-n\varphi) + i \operatorname{sen}(-n\varphi)) \tag{3.18}$$

Las series de Taylor para $\cos\varphi$ y $\operatorname{sen}\varphi$ se pueden escribir en forma compleja,

$$\begin{aligned} \cos\varphi &= 1 - \frac{\varphi^2}{2!} + \frac{\varphi^4}{4!} - \cdots = 1 + \frac{(i\varphi)^2}{2!} + \frac{(i\varphi)^4}{4!} + \cdots \\ i \operatorname{sen}\varphi &= i\left(\varphi - \frac{\varphi^3}{3!} + \frac{\varphi^5}{5!} - \cdots\right) = i\varphi + \frac{(i\varphi)^3}{3!} + \frac{(i\varphi)^5}{5!} + \cdots \end{aligned}$$

Por tanto,

$$\cos\varphi + i \operatorname{sen}\varphi = e^{i\varphi} \tag{3.19}$$

y de la representación polar se obtiene la **representación exponencial** para un número complejo $z$:

$$z = \rho (\cos\varphi + i \operatorname{sen}\varphi) = \rho e^{i\varphi} \tag{3.20}$$

Esta representación es conveniente para la multiplicación y conjugación. Si

$$z = x + iy = \rho (\cos\varphi + i \operatorname{sen}\varphi) = \rho e^{i\varphi}$$

entonces,

$$z^* = x - iy = \rho (\cos\varphi - i \operatorname{sen}\varphi) = \rho (\cos(-\varphi) + i \operatorname{sen}(-\varphi)) = \rho e^{-i\varphi} \tag{3.21}$$

Además, se tiene

$$z_1 \cdot z_2 = \rho_1\rho_2 (\cos(\varphi_1 + \varphi_2) + i \operatorname{sen}(\varphi_1 + \varphi_2)) = \rho_1\rho_2 e^{i(\varphi_1+\varphi_2)} \tag{3.22}$$

$$(z_1 z_2)^* = z_1^* z_2^*, \quad |z|^2 = zz^* = \rho^2 \tag{3.23}$$

## 3.2. Funciones complejas básicas

Una magnitud compleja $z = x + iy$, en la cual $x$ y $y$ están consideradas como variables reales, se llama **variable compleja**. Por función de una variable compleja $z$ definida sobre un conjunto $D$ se entiende una correspondencia (regla) que asigna a cada $z$ en $D$ un número complejo $w = u(x, y) + iv(x, y)$, en donde $u(x, y)$ y $v(x, y)$ son dos funciones reales, dependiendo cada una de las dos variables reales $x$ y $y$. Entonces, se escribe $w = f(z)$. El conjunto $D$ es el **dominio de definición** de $f(z)$. La función $w = f(z)$ se llama **univaluada** en una región $D$ si a cada número complejo $z$ de esta región le corresponde un valor único de $w$. Se dice que $w = f(z)$ es una función **multivaluada** si a cada $z$ le corresponde más de un valor de $w$.

**Ejemplo 3.1.** *a) La función lineal $w = az$, es decir, $u = ax$ y $v = ay$, en donde $a$ es una constante real, es univaluada en todo el plano complejo $z$.*

*b) La función $w = z^2 = (x^2 - y^2) + i2xy$, es decir, $u = x^2 - y^2$ y $v = 2xy$, es también univaluada en todo el plano $z$.*

*c) La función $w = zz^*$, en donde $u = x^2 + y^2$ y $v = 0$ es univaluada.*

*d) Es fácil ver que la ecuación $w^2 = z$ tiene la solución $w = \sqrt{z}$ $= \sqrt{\rho}\left(\cos\frac{\varphi+2\pi k}{2} + i\operatorname{sen}\frac{\varphi+2\pi k}{2}\right)$, $k = 0, 1$, en donde la variable compleja es $z = \rho(\cos\varphi + i\operatorname{sen}\varphi)$. Por lo tanto, la función $w = \sqrt{z}$, que se define como la solución de $w^2 = z$, es doble valuada: $w_1 = \sqrt{\rho}\left(\cos\frac{\varphi}{2} + i\operatorname{sen}\frac{\varphi}{2}\right)$, $w_2 = \sqrt{\rho}(\cos\left(\frac{\varphi}{2} + \pi\right) + i\operatorname{sen}\left(\frac{\varphi}{2} + \pi\right))$.*

En esta sección se consideran las funciones complejas básicas y sus propiedades. Al considerar funciones complejas, es necesario distinguir entre los puntos internos y los de frontera de un dominio. Se dice que un punto de dominio $D$ es **punto interno** si existe un círculo de un radio no igual a cero, con el centro en este punto y tal que contiene en su interior únicamente los puntos del dominio $D$. Los puntos de la frontera de $D$ no son interiores, dado que cualquier círculo alrededor de un punto de frontera necesariamente contiene tanto puntos de $D$ como puntos que no pertenecen al dominio $D$. La frontera de un dominio $D$ puede ser constituida por los puntos de $D$ o los que no pertenecen a $D$. Un dominio constituido únicamente por puntos interiores se llama abierto. En este caso la frontera no pertenece a éste. Un ejemplo de **dominio abierto** es $|z| < R$. Cuando los puntos de frontera están incluidos en el dominio, éste se llama **cerrado**. Un ejemplo de región cerrada (de dominio cerrado) es $|z| \leq R$. Cuando cada punto de un dominio se encuentra a una distancia finita del centro, se dice que el **dominio es limitado** o **finito**. Todos los puntos de un dominio limitado se encuentran adentro de un círculo $|z| = R$, si el radio $R$ es escogido suficientemente grande. El dominio que no es limitado se llama no limitado. Un ejemplo, $|z| \geq 1$.

### 3.2.1. Funciones algebraicas de una variable compleja

Sea $z = \rho(\cos\varphi + i\,\mathrm{sen}\,\varphi) = \rho e^{i\varphi}$. La función $w = z^n$, potencia entera $n = 0, \pm1, \pm2, \ldots$ de $z$, es definida mediante la relación (véase la sección anterior):

$$w = z^n = \rho^n(\cos n\varphi + i\,\mathrm{sen}\,n\varphi) = \rho^n e^{in\varphi} \tag{3.24}$$

Entonces, representando $w = R(\cos\theta + i\,\mathrm{sen}\,\theta)$, se tiene $R = \rho^n$ y $\theta = \arg w = n\varphi = n\arg z$. Nótese que la función $w = z^n$ es univaluada, es decir, a cada valor de $z$ le corresponde un valor único de $w$.

Para extender el concepto de la función de potencias enteras a la de las potencias racionales, $a = m/n$, basta demostrarlo para la función

$$w = z^{1/n} = \sqrt[n]{z} \tag{3.25}$$

y después elevar el resultado a una potencia entera $m$. La función $w = z^{1/n}$ se define como la solución de la ecuación

$$w^n = z \tag{3.26}$$

En representación polar, $w = R(\cos\theta + i\,\mathrm{sen}\,\theta)$ y $z = \rho(\cos\varphi + i\,\mathrm{sen}\,\varphi)$, de la última ecuación con base en la fórmula de De Moivre se tiene

$$\begin{aligned} R^n(\cos n\theta + i\,\mathrm{sen}\,n\theta) &= \rho(\cos\varphi + i\,\mathrm{sen}\,\varphi) \\ &= \rho(\cos(\varphi + 2\pi k) + i\,\mathrm{sen}(\varphi + 2\pi k)) \end{aligned}$$

en donde $\varphi$ es el argumento principal de $z$. Al igualar los módulos y los argumentos se tiene

$$R^n = \rho, \quad n\theta = \varphi + 2\pi k$$

en donde $k$ es un entero. Entonces,

$$R = \rho^{1/n}, \quad \theta = \frac{\varphi + 2\pi k}{n}$$

En consecuencia, la función $w = z^{1/n} = \sqrt[n]{z}$ para $z \neq 0$ es multivaluada (de valores múltiples) y tiene $n$ valores distintos

$$w = z^{1/n} = \rho^{1/n}\left(\cos\left(\frac{\varphi + 2\pi k}{n}\right) + i\,\mathrm{sen}\left(\frac{\varphi + 2\pi k}{n}\right)\right), \quad k = 0, 1, \ldots, n-1 \tag{3.27}$$

para cada valor de la variable $z$. El valor de $w$ obtenido para el valor principal de $\arg z = \varphi$ y $k = 0$ se llama valor principal de $w = z^{1/n}$. De la ecuación (3.27) se obtiene que los valores de $w = z^{1/n}$ son los $n$ vértices de un polígono inscrito en la circunferencia de radio $\rho^{1/n}$ (véase la figura 3.2). Cada uno de estos $n$ valores se llama $n$-ésima raíz de $z$.

**Ejemplo 3.2.** *Dado que* $1 = \cos(2\pi k) + i\,\mathrm{sen}(2\pi k)$, *la raíz cúbica de la unidad* $\sqrt[3]{1}$ *tiene tres valores,*

$$\begin{aligned} w_k &= \cos\frac{2\pi k}{3} + i\,\mathrm{sen}\,\frac{2\pi k}{3}, \quad k = 0, 1, 2 \\ w_1 &= 1, \quad w_2 = -1/2 + i\sqrt{3}/2, \quad w_3 = -1/2 - i\sqrt{3}/2 \end{aligned}$$

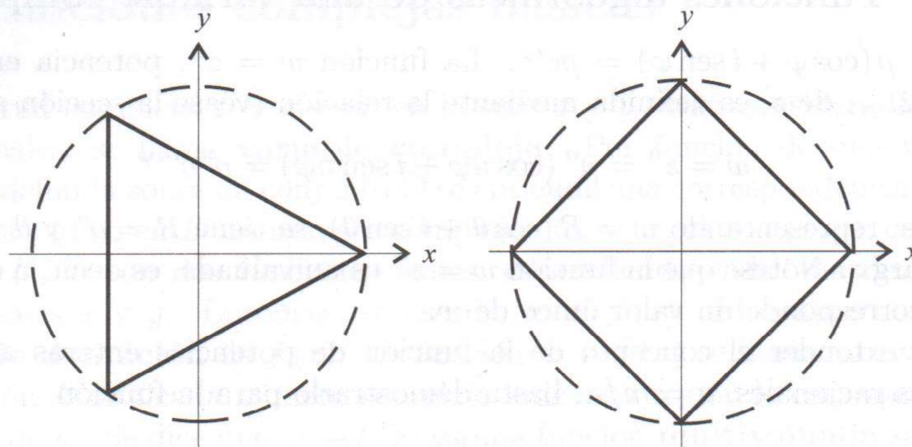



Figura 3.2. Raíz cúbica y cuarta de 1

**Ejemplo 3.3.** *Dado que* $1+i=\sqrt{2}\,(\cos(\pi/4)+i\,\mathrm{sen}\,(\pi/4))$, *la raíz cúbica* $\sqrt[3]{1+i}$ *tiene tres valores*

$$w_k=\sqrt[3]{2}\left(\cos\frac{\pi/4+2\pi k}{3}+i\,\mathrm{sen}\,\frac{\pi/4+2\pi k}{3}\right),\quad k=0,1,2$$

En los párrafos anteriores definimos las operaciones de adición, sustracción, multiplicación, división y, además, las funciones de potencia entera y potencia racional de una variable compleja $z$. Esto es suficiente para determinar valores de funciones algebraicas como

$$w=\frac{a_0z^m+a_1z^{m-1}+\cdots+a}{b_0z^n+b_1z^{n-1}+\cdots+b}$$

en donde las potencias $m$ y $n$ pueden ser números enteros o fracciones.

### 3.2.2. Funciones trascendentales elementales

Las funciones que no pueden representarse como algebraicas (por ejemplo, trigonométricas, logarítmicas) se llaman trascendentales. Las homólogas complejas de las funciones trascendentales elementales reales $e^x$, $\ln x$, $x^a$, $\mathrm{senh}\,x$, $\cosh x$, $\tanh x$, etcétera, no son funciones algebraicas y, por tanto, requieren su adecuada definición. Una definición de funciones complejas sería útil si preserva las propiedades esenciales de las funciones reales, por ejemplo la ley de exponentes, $e^xe^y=e^{x+y}$.

#### Exponencial

La exponencial de una variable compleja $z=x+iy$ se define como sigue:

$$w=e^z=e^{x+iy}=e^x\,(\cos y+i\,\mathrm{sen}\,y) \tag{3.28}$$

ésta es función univaluada. Nótese que para $x = 0$ de la ecuación (3.28) se tiene

$$e^{iy} = \cos y + i \operatorname{sen} y \tag{3.29}$$

que coincide con la fórmula (3.19). Cambiando $y$ por $-y$ en la ecuación (3.29) se tiene

$$e^{-iy} = \cos y - i \operatorname{sen} y$$

Sumando y sustrayendo las fórmulas anteriores, se obtienen las fórmulas de Euler:

$$\cos y = \frac{1}{2}\left(e^{iy} + e^{-iy}\right), \quad \operatorname{sen} y = \frac{1}{2i}\left(e^{iy} - e^{-iy}\right) \tag{3.30}$$

De la definición, ecuación (3.28), aplicando las fórmulas trigonométricas, con facilidad se encuentra

$$\begin{aligned} e^{z_1} e^{z_2} &= e^{x_1} (\cos y_1 + i \operatorname{sen} y_1)\, e^{x_2} (\cos y_2 + i \operatorname{sen} y_2) \\ &= e^{x_1 + x_2} (\cos (y_1 + y_2) + i \operatorname{sen} (y_1 + y_2)) = e^{z_1 + z_2} \end{aligned} \tag{3.31}$$

Dado que $e^{i2\pi k} = 1$, se encuentra que

$$e^{z + i2\pi k} = e^z e^{i2\pi k} = e^z, \quad k = 0, \pm 1, \pm 2, \ldots$$

pues la función $e^z$ es periódica con el periodo imaginario $i2\pi$.

**Ejemplo 3.4.** *De la ecuación (3.28) se obtiene que*

$$e^{1-i} = e^1 e^{-i} = e\,(\cos 1 - i \operatorname{sen} 1)$$

### Funciones trigonométricas

Las funciones trigonométricas de variable compleja $z$ se definen a través de exponenciales de la siguiente manera:

$$\cos z = \frac{1}{2}\left(e^{iz} + e^{-iz}\right), \quad \operatorname{sen} z = \frac{1}{2i}\left(e^{iz} - e^{-iz}\right) \tag{3.32}$$

$$\tan z = \frac{\operatorname{sen} z}{\cos z}, \quad \cot z = \frac{\cos z}{\operatorname{sen} z}, \quad \sec z = \frac{1}{\cos z}, \quad \csc z = \frac{1}{\operatorname{sen} z} \tag{3.33}$$

Éstas son funciones univaluadas. Usando las definiciones, con facilidad se establece que todas las fórmulas de trigonometría analítica de variable real son válidas para las funciones trigonométricas de variable compleja. Por ejemplo,

$$\cos(-z) = \cos z,\ \operatorname{sen}(-z) = -\operatorname{sen} z,\ \cos(z + 2\pi k) = \cos z \tag{3.34}$$

$$\operatorname{sen}^2 z + \cos^2 z = 1, \quad \operatorname{sen}(z_1 \pm z_2) = \operatorname{sen} z_1 \cos z_2 \pm \cos z_1 \operatorname{sen} z_2 \tag{3.35}$$

De la ecuación (3.32) se ve que las fórmulas de Euler son válidas para los valores complejos,

$$e^{iz} = \cos z + i \operatorname{sen} z, \quad e^{-iz} = \cos z - i \operatorname{sen} z \tag{3.36}$$

**Ejemplo 3.5.** *Para calcular* $\operatorname{sen}(1-i)$ *podemos seguir dos caminos. Primero, según la definición,*

$$\begin{aligned}\operatorname{sen}(1-i) &= \frac{e^{i(1-i)}-e^{-i(1-i)}}{2i}=\frac{e^{i+1}-e^{-i-1}}{2i}\\ &= \frac{1}{2i}\left[e^{1}\left(\cos 1+i\operatorname{sen}1\right)-e^{-1}\left(\cos 1-i\operatorname{sen}1\right)\right]\\ &= \frac{e^{1}-e^{-1}}{2i}\cos 1+\frac{e^{1}+e^{-1}}{2}\operatorname{sen}1\end{aligned}$$

*Por otra parte, aplicando la fórmula trigonométrica, se tiene*

$$\operatorname{sen}(1-i)=\operatorname{sen}1\cos i-\cos 1\operatorname{sen}i$$

*Sustituyendo*

$$\cos i=\left(e^{-1}+e^{1}\right)/2 \quad y \quad \operatorname{sen}i=\left(e^{-1}-e^{1}\right)/2i$$

*en la fórmula anterior, se tiene el resultado que obtuvimos de la definición.*

### Funciones hiperbólicas

Las funciones hiperbólicas de una variable compleja $z$ se definen mediante las fórmulas:

$$\cosh z = \frac{1}{2}\left(e^{z}+e^{-z}\right), \quad \operatorname{senh}z=\frac{1}{2}\left(e^{z}-e^{-z}\right) \tag{3.37}$$

$$\tanh z = \frac{\operatorname{senh}z}{\cosh z}, \quad \coth z=\frac{\cosh z}{\operatorname{senh}z} \tag{3.38}$$

$$\operatorname{sech}z = \frac{1}{\cosh z}, \quad \operatorname{csch}z=\frac{1}{\operatorname{senh}z} \tag{3.39}$$

Éstas son univaluadas. De las ecuaciones (3.32) y (3.37) se obtiene que

$$\cosh(iz) = \cos z, \quad \operatorname{senh}(iz)=i\operatorname{sen}z \tag{3.40}$$

$$\cos(iz) = \cosh z, \quad \operatorname{sen}(iz)=i\operatorname{senh}z \tag{3.41}$$

### Logaritmo natural

El logaritmo de un número complejo $z$ se define de manera similar que el logaritmo de una variable real, pues el logaritmo de base natural

$$w=\ln z \tag{3.42}$$

de un número complejo $z$ es la solución de la ecuación

$$e^{w}=z \tag{3.43}$$

Puesto que $w=u+iv$ y $z=\rho\left(\cos\varphi+i\operatorname{sen}\varphi\right)$, de la ecuación (3.43) se tiene

$$e^{u}\left(\cos v+i\operatorname{sen}v\right)=\rho\left(\cos\varphi+i\operatorname{sen}\varphi\right)$$

Esto es,

$$e^u = \rho, \quad v = \varphi + 2\pi k, \quad k = 0, \pm 1, \pm 2, \ldots$$

Dado que $u$ y $v$ son reales, de la ecuación anterior concluimos que $u = \ln \rho$, en donde $\ln \rho$ es el logaritmo de la variable real. Finalmente, se obtiene

$$w = \ln z = u + iv = \ln \rho + i\,(\varphi + 2\pi k) = \frac{1}{2} \ln \left(x^2 + y^2\right) + i \arctan \frac{y}{x} \qquad (3.44)$$

en donde $\arctan \frac{y}{x} = \varphi + 2\pi k, \quad k = 0, \pm 1, \pm 2, \ldots$ Entonces, el logaritmo natural de un número complejo tiene una infinidad de valores; es decir, el logaritmo de una variable compleja es una función infinitamente valuada. Siendo $k = 0$ y $0 \leq \varphi < 2\pi$ el argumento principal, se obtiene la función univaluada

$$\text{Ln } z = \ln \rho + i\varphi \qquad (3.45)$$

que recibe el nombre de valor principal de $\ln z$.

**Ejemplo 3.6.** *Dado que* $1 + i = \sqrt{2}\,(\cos(\pi/4) + i \operatorname{sen}(\pi/4))$, *se tiene que*

$$\ln(1+i) = \ln\sqrt{2} + i \arctan 1 = \frac{1}{2} \ln 2 + i\left(\frac{\pi}{4} + 2\pi k\right)$$

*en donde* $k = 0, \pm 1, \pm 2, \ldots$

## Potencia compleja

Las potencias enteras son univaluadas; las potencias racionales, multivaluadas. La definición de logaritmo, ecuación (3.44) nos sirve para definir las potencias complejas e irracionales por la fórmula

$$z^c \equiv e^{c \ln z} \qquad (3.46)$$

en donde $c$ es un número complejo, en particular entero, racional o irracional. Basándose en las ecuaciones (3.42) y (3.43) se ve que esta definición es equivalente a

$$c \ln z = \ln z^c \qquad (3.47)$$

Por el hecho de que el logaritmo es una función infinitamente valuada, entonces, la función $z^c$, por lo general, es también infinitamente valuada. Sin embargo, en plena concordancia con los resultados para potencias enteras y racionales, la ecuación (3.46) conduce a una función univaluada o multivaluada, cuando la potencia $c$ es un número entero o racional, respectivamente.

**Ejemplo 3.7.** *Siguiendo la definición, ecuación (3.46), se tiene que*

$$2^i = e^{i \ln 2} = e^{i(\ln 2 + i2\pi k)} = e^{i \ln 2 - 2\pi k} = e^{-2\pi k}\left(\cos(\ln 2) + i \operatorname{sen}(\ln 2)\right)$$

**Ejemplo 3.8.** *De manera similar se tiene que*

$$\begin{aligned}
(1+i)^{(2+i)} &= e^{(2+i)\ln(1+i)} \\
&= e^{(2+i)\left[\frac{1}{2}\ln 2 + i(\pi/4 + 2\pi k)\right]} = e^{\ln 2 - \pi/4 - 2\pi k}\, e^{i\left[\frac{1}{2}\ln 2 + \pi/2 + 4\pi k\right]} \\
&= e^{\ln 2 - \pi/4 - 2\pi k}\left[\cos\left(\frac{1}{2}\ln 2 + \frac{\pi}{2}\right) + i \operatorname{sen}\left(\frac{1}{2}\ln 2 + \frac{\pi}{2}\right)\right]
\end{aligned}$$

# 3.3. Funciones analíticas

## 3.3.1. Función analítica. Teorema de Cauchy-Riemann. Ecuación de Laplace

En la presente sección se definen las funciones analíticas. Siendo una clase particular de las funciones complejas, éstas son de gran importancia debido a sus propiedades especiales, que se estudian a continuación.

Se dice que un punto $z = x + iy$ tiende al punto $z_0 = x_0 + iy_0$, si $x \to x_0$ y $y \to y_0$. Sea $f(z)$ una función univaluada en una vecindad del punto $z_0$. Por la **vecindad** de $z_0$ se entiende el conjunto de todos los puntos de un dominio circular, $|z - z_0| < \delta$, suficientemente pequeño con el centro en $z_0$. Mientras $z \to z_0$, la función $f(z)$ tiende a un valor $w_0$ que puede, por lo general, depender de la trayectoria de acercamiento al punto $z_0$. Entonces, se dice que $w_0$ es el límite de $f(z)$ al aproximarse $z \to z_0$ y se escribe

$$w_0 = \lim_{z \to z_0} f(z)$$

En particular, si $w_0 = f(z_0)$, entonces, se dice que $f(z)$ es **continua en el punto** $z_0$. Se dice que $f(z)$ es **continua en un dominio** si es continua en cada punto de ese dominio. Es fácil comprobar que $f(z) = u(x, y) + iv(x, y)$ es continua en $z_0 = x_0 + iy_0$ si y sólo si su parte real $u(x, y)$ e imaginaria $v(x, y)$ son funciones continuas en $(x_0, y_0)$.

Suponemos que la función $w = f(z) = u(x, y) + iv(x, y)$ es continua en un dominio del plano complejo $z$. Las magnitudes complejas $z$ y $w$ pueden representarse como puntos en dos planos complejos distintos, llamados plano $z$ y plano $w$. Sean $z_0 = x_0 + iy_0$ y $z = z_0 + \Delta z$ dos puntos cercanos en el plano complejo $z$, siendo $\Delta z = \Delta x + i\Delta y$. Los puntos correspondientes en el plano complejo $w$ sean $w_0 = u_0 + iv_0$ y $w = w_0 + \Delta w$, en donde

$$\Delta w = f(z_0 + \Delta z) - f(z_0) = \Delta u + i\Delta v$$

y $u_0 = u(x_0, y_0)$, $v_0 = v(x_0, y_0)$, $\Delta u = u - u_0$, $\Delta v = v - v_0$. Se dice que una función $w = f(z)$ es **diferenciable en el punto** $z_0$ si el límite

$$f'(z_0) = \lim_{\Delta z \to 0} \frac{\Delta w}{\Delta z} = \lim_{\Delta z \to 0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z} \tag{3.48}$$

existe y es único, es decir, el límite es independiente de la trayectoria que conecta los puntos $z$ y $z_0$, mientras $\Delta z$ tiende a cero. Este límite se conoce como **derivada** de $f(z)$ en el punto $z_0$.

**Ejemplo 3.9.** *Considérese la función $w = zz^* = |z|^2$. Se tiene*

$$\begin{aligned} \Delta w &= (z + \Delta z)(z^* + \Delta z^*) - zz^* = z^* \Delta z + z(\Delta z)^* + \Delta z \Delta z^* \\ \frac{\Delta w}{\Delta z} &= z^* + z\frac{(\Delta z)^*}{\Delta z} + \Delta z^* \end{aligned} \tag{3.49}$$

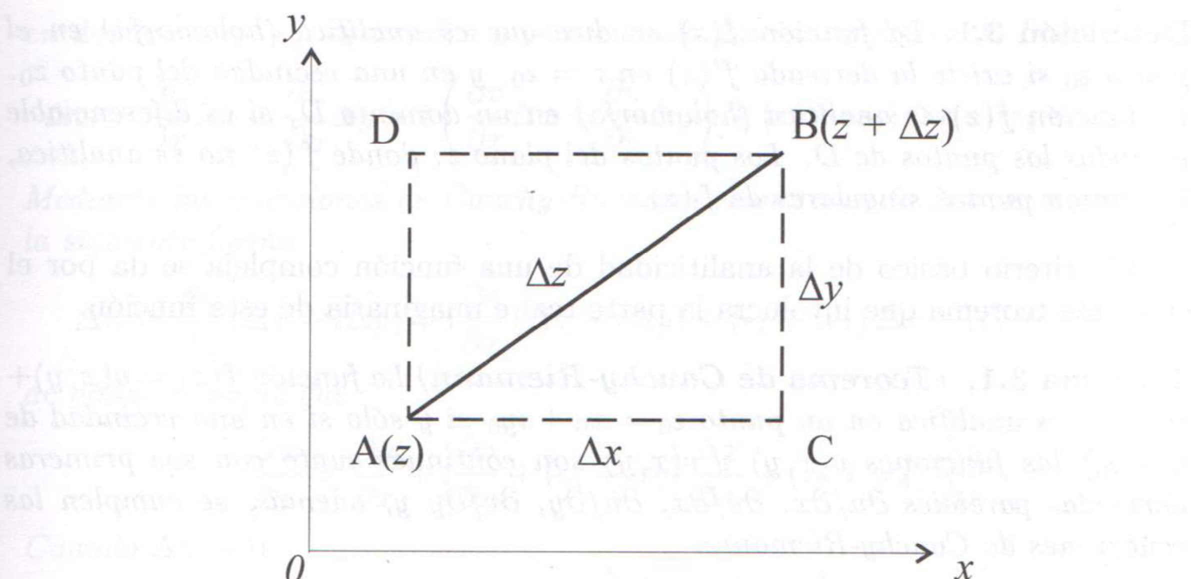



Figura 3.3. Dos trayectorias posibles

*A continuación se demuestra que este cociente, por lo general, no tiene un límite único cuando $\Delta z \to 0$ a lo largo de diferentes caminos. Sea $z = x + iy$ y $\Delta z = \Delta x + i\Delta y$. Escogiendo el camino BCA (véase la figura 3.3), se tiene:*

$$\begin{aligned}\lim_{\Delta z\to 0}\frac{\Delta w}{\Delta z} &= \lim_{\Delta x\to 0}\lim_{\Delta y\to 0}\frac{\Delta w}{\Delta z}\\ &= \lim_{\Delta x\to 0}\lim_{\Delta y\to 0}\left(z^* + z\frac{\Delta x - i\Delta y}{\Delta x + i\Delta y} + (\Delta x - i\Delta y)\right) = 2x\end{aligned}$$

*Para la trayectoria BDA:*

$$\begin{aligned}\lim_{\Delta z\to 0}\frac{\Delta w}{\Delta z} &= \lim_{\Delta y\to 0}\lim_{\Delta x\to 0}\frac{\Delta w}{\Delta z}\\ &= \lim_{\Delta y\to 0}\lim_{\Delta x\to 0}\left(z^* + z\frac{\Delta x - i\Delta y}{\Delta x + i\Delta y} + (\Delta x - i\Delta y)\right) = -i2y\end{aligned}$$

*Como los límites son diferentes, la función no tiene derivada en cualquier punto del plano $z$, salvo el punto $z = 0$.*

**Ejemplo 3.10.** *La función $w = f(z) = z^2$ es diferenciable para toda $z$ y su derivada es $f'(z) = 2z$ porque*

$$\lim_{\Delta z\to 0}\frac{f(z+\Delta z) - f(z)}{\Delta z} = \lim_{\Delta z\to 0}\frac{(z+\Delta z)^2 - z^2}{\Delta z} = \lim_{\Delta z\to 0}(2z + \Delta z) = 2z$$

*Como éste es un límite único, $w = z^2$ sí tiene la derivada $f'(z) = 2z$.*

Es fácil demostrar que todas las reglas conocidas del Cálculo Diferencial real, como las reglas para derivar una constante, las potencias enteras de $z$, sumas, productos y cocientes de funciones diferenciables, y la regla de la cadena, son válidas para las funciones complejas diferenciables.

**Definición 3.1.** *La función $f(z)$ se dice que es analítica (holomorfa) en el punto $z_0$ si existe la derivada $f'(z)$ en $z = z_0$ y en una vecindad del punto $z_0$. La función $f(z)$ es analítica (holomorfa) en un dominio $D$, si es diferenciable en todos los puntos de $D$. Los puntos del plano $z$, donde $f(z)$ no es analítica, se llaman puntos singulares de $f(z)$.*

El criterio básico de la analiticidad de una función compleja se da por el siguiente teorema que involucra la parte real e imaginaria de esta función.

**Teorema 3.1.** ***(Teorema de Cauchy-Riemann)*** *La función $f(z) = u(x,y) + iv(x,y)$ es analítica en un punto $z_0 = x_0 + iy_0$ si y sólo si en una vecindad de $(x_0, y_0)$ las funciones $u(x,y)$ y $v(x,y)$ son continuas junto con sus primeras derivadas parciales $\partial u/\partial x$, $\partial v/\partial x$, $\partial u/\partial y$, $\partial v/\partial y$ y, además, se cumplen las ecuaciones de Cauchy-Riemann:*

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y} \tag{3.50}$$

**Demostración.** *Condiciones necesarias. Suponemos que $f(z)$ es analítica en $z$. Dado que $\Delta z = \Delta x + i\Delta y$ y $\Delta w = \Delta u(x,y) + i\Delta v(x,y)$, se tiene que el límite*

$$f'(z) = \lim_{\Delta z \to 0} \frac{\Delta u + i\Delta v}{\Delta z} = \lim_{\Delta x, \Delta y \to 0} \frac{\Delta u + i\Delta v}{\Delta x + i\Delta y}$$

*es único. Para evaluar el límite, primero hacemos tender $\Delta z \to 0$ de la siguiente manera:*

$$f'(z) = \lim_{\Delta x \to 0} \lim_{\Delta y \to 0} \frac{\Delta u + i\Delta v}{\Delta x + i\Delta y} = \lim_{\Delta x \to 0} \left( \frac{\hat{\Delta} u}{\Delta x} + i\frac{\hat{\Delta} v}{\Delta x} \right) = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}$$

*en donde $\hat{\Delta} u = u(x + \Delta x, y) - u(x,y)$, $\hat{\Delta} v = v(x + \Delta x, y) - v(x,y)$. Luego, tendiendo primero $\Delta x \to 0$ y, después, $\Delta y \to 0$, se obtiene*

$$f'(z) = \lim_{\Delta y \to 0} \lim_{\Delta x \to 0} \frac{\Delta u + i\Delta v}{\Delta x + i\Delta y} = \lim_{\Delta y \to 0} \left( \frac{\check{\Delta} u}{i\Delta y} + \frac{\check{\Delta} v}{\Delta y} \right) = \frac{\partial v}{\partial y} - i\frac{\partial u}{\partial y}$$

*en donde $\check{\Delta} u = u(x, y + \Delta y) - u(x,y)$, $\check{\Delta} v = v(x, y + \Delta y) - v(x,y)$. Entonces, para que estos límites tengan los mismos valores en el punto $z$, deben satisfacerse las ecuaciones*

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}$$

*que reciben el nombre de ecuaciones de Cauchy-Riemann.*

*Condiciones suficientes. Suponemos que se cumplen las ecuaciones de Cauchy-Riemann y, además, $u(x,y)$ y $v(x,y)$ son diferenciables, es decir,*

$$\Delta u = u(x + \Delta x, y + \Delta y) - u(x,y) = \frac{\partial u}{\partial x}\Delta x + \frac{\partial u}{\partial y}\Delta y + \epsilon_1 \Delta x + \epsilon_2 \Delta y$$

$$\Delta v = v(x + \Delta x, y + \Delta y) - v(x,y) = \frac{\partial v}{\partial x}\Delta x + \frac{\partial v}{\partial y}\Delta y + \epsilon_3 \Delta x + \epsilon_4 \Delta y$$

*en donde $\epsilon_1$, $\epsilon_2$, $\epsilon_3$, $\epsilon_4$ tienden a cero cuando $\Delta x \to 0$ y $\Delta y \to 0$. Luego,*

$$\Delta w = \frac{\partial u}{\partial x}\Delta x + \frac{\partial u}{\partial y}\Delta y + i\left(\frac{\partial v}{\partial x}\Delta x + \frac{\partial v}{\partial y}\Delta y\right) + (\epsilon_1 + i\epsilon_3)\,\Delta x + (\epsilon_2 + i\epsilon_4)\,\Delta y$$

*Mediante las ecuaciones de Cauchy-Riemann el incremento $\Delta w$ se reescribe de la siguiente forma:*

$$\Delta w = \frac{\partial u}{\partial x}(\Delta x + i\Delta y) + i\frac{\partial v}{\partial x}(\Delta x + i\Delta y) + (\epsilon_1 + i\epsilon_3)\,\Delta x + (\epsilon_2 + i\epsilon_4)\,\Delta y$$

*de donde se sigue que*

$$\frac{\Delta w}{\Delta z} = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} + (\epsilon_1 + i\epsilon_3)\,\frac{\Delta x}{\Delta z} + (\epsilon_2 + i\epsilon_4)\,\frac{\Delta y}{\Delta z}$$

*Cuando $\Delta z \to 0$,*

$$f'(z) = \lim_{\Delta z \to 0}\frac{\Delta w}{\Delta z} = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} \tag{3.51}$$

*ya que las expresiones*

$$\left|(\epsilon_1 + i\epsilon_3)\,\frac{\Delta x}{\Delta z}\right| \le |\epsilon_1 + i\epsilon_3|\,, \qquad \left|(\epsilon_2 + i\epsilon_4)\,\frac{\Delta y}{\Delta z}\right| \le |\epsilon_2 + i\epsilon_4|$$

*tienden a cero cuando $\Delta x \to 0$ y $\Delta y \to 0$. Entonces, en cada punto de una vecindad de $z$ existe la derivada de $f(z)$, ecuación (3.51), y queda completa la demostración.* ■

Nótese que las reglas de diferenciación de funciones compuestas e inversas reales se extienden para las funciones complejas analíticas:

$$\frac{dF(w)}{dz} = \frac{dF(w)}{dw}\frac{dw}{dz} \tag{3.52}$$

$$\frac{dz}{dw} = \frac{1}{dw/dz} \quad \text{si} \quad \frac{dw}{dz} \neq 0 \tag{3.53}$$

**Ejemplo 3.11.** *La función $w = f(z) = z^2 = (x^2 - y^2) + i2xy$ es analítica en el plano $z$, ya que se cumplen las ecuaciones de Cauchy-Riemann:*

$$\frac{\partial u}{\partial x} = 2x, \quad \frac{\partial v}{\partial y} = 2x, \quad \frac{\partial v}{\partial x} = 2y, \quad \frac{\partial u}{\partial y} = -2y$$

*Luego,*

$$\frac{\partial(z^2)}{\partial z} = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} = 2x + i2y = 2(x + iy) = 2z$$

**Ejemplo 3.12.** *La exponencial $w \equiv u + iv = e^z = e^x(\cos y + i\,\mathrm{sen}\,y)$ es función analítica en todo el plano $z$, porque*

$$\frac{\partial u}{\partial x} = e^x\cos y, \quad \frac{\partial v}{\partial y} = e^x\cos y, \quad \frac{\partial v}{\partial x} = e^x\,\mathrm{sen}\,y, \quad \frac{\partial u}{\partial y} = -e^x\,\mathrm{sen}\,y$$

*Además,*

$$\frac{d(e^z)}{dz} = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} = e^x\cos y + ie^x\,\mathrm{sen}\,y = e^z$$

**Ejemplo 3.13.** *La función $w = \ln z$ es multivaluada. Sin embargo, cada rama de ésta, determinada por un valor fijo del número $k$, es univaluada y, verificando las condiciones de Cauchy-Riemann, se obtiene que es función analítica, excepto el punto $z = 0$. Fijando un número $k$ y usando la regla de diferenciación para la función inversa $z = e^w$, se obtiene la derivada del logaritmo complejo,*

$$\frac{dw}{dz} = \frac{d(\ln z)}{dz} = \frac{1}{dz/dw} = \frac{1}{z} \quad si \quad z \neq 0$$

*que es la misma para diferentes ramas de esta función. El punto $z = 0$ es singular de la función $w = \ln z$, porque no existe la derivada en este punto.*

Existe una relación importante entre las funciones analíticas y la ecuación de Laplace en dos variables. Esta relación es consecuencia de las ecuaciones de Cauchy-Riemann y el hecho, que se probará posteriormente, de que la derivada de una función analítica $f(z) = u(x, y) + iv(x, y)$ también es analítica. Por este hecho, $u(x, y)$ y $v(x, y)$ tienen derivadas parciales continuas de todos los órdenes y, en particular, las segundas derivadas mixtas de estas funciones son iguales:

$$\frac{\partial^2 u}{\partial x \partial y} = \frac{\partial^2 u}{\partial y \partial x}, \quad \frac{\partial^2 v}{\partial x \partial y} = \frac{\partial^2 v}{\partial y \partial x}$$

Al derivar las ecuaciones de Cauchy-Riemann y utilizando las igualdades anteriores, se obtiene

$$\begin{aligned} \frac{\partial^2 u}{\partial x^2} &= \frac{\partial^2 v}{\partial x \partial y}, \quad \frac{\partial^2 u}{\partial y^2} = -\frac{\partial^2 v}{\partial x \partial y} \\ \frac{\partial^2 v}{\partial x^2} &= -\frac{\partial^2 u}{\partial x \partial y}, \quad \frac{\partial^2 v}{\partial y^2} = \frac{\partial^2 u}{\partial x \partial y} \end{aligned}$$

Esto conduce al siguiente teorema.

**Teorema 3.2.** *Las partes real e imaginaria de una función compleja $f(z) = u(x, y) + iv(x, y)$, que es analítica en un dominio $D$, son soluciones de la ecuación de Laplace*

$$\nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad \nabla^2 v = \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0 \tag{3.54}$$

*en $D$ y tienen segundas derivadas parciales continuas en éste.*

Una solución de la ecuación de Laplace que tenga derivadas parciales de segundo orden continuas se conoce como **función armónica**. Por tanto, las partes real e imaginaria de una función analítica son funciones armónicas.

Si dos funciones armónicas, $u(x, y)$ y $v(x, y)$, satisfacen las ecuaciones de Cauchy-Riemann en un dominio $D$, es decir, si $u(x, y)$ y $v(x, y)$ son las partes real e imaginaria de una función analítica en $D$, entonces se dice que $v(x, y)$ es una función armónica conjugada de $u(x, y)$ en $D$. Nótese que, conociendo la parte real (imaginaria) de una función analítica, la parte imaginaria (real) de ésta se recupera resolviendo las ecuaciones de Cauchy-Riemann, es decir, una función analítica es completamente definida por su parte real o imaginaria.

**Ejemplo 3.14.** *Para encontrar la parte imaginaria $v(x, y)$ de una función analítica, que tiene como la parte real la función armónica $u(x, y) = x^2 - y^2$, se procede de la siguiente manera. Se tiene*

$$\frac{\partial u}{\partial x} = 2x = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -2y = -\frac{\partial v}{\partial x}$$

*Al integrar la primera ecuación con respecto a $y$, se obtiene*

$$v(x, y) = 2xy + h(x)$$

*en donde $h(x)$ depende sólo de $x$. Sustituyendo esta solución en la segunda ecuación de Cauchy-Riemann, se tiene $h'(x) = 0$ y por tanto $h(x) = C$ es una constante real. Entonces, la función analítica más general con la parte real $u(x, y) = x^2 - y^2$ es $f(z) = x^2 - y^2 + i(2xy + C) = z^2 + iC$.*

### 3.3.2. Integración de funciones complejas. Teorema integral de Cauchy

La importancia de las integrales complejas se debe principalmente a dos razones. Primera, la integración compleja proporciona demostraciones de algunas propiedades de las funciones analíticas, las cuales serían difíciles de comprobar sin utilizar el método de integración compleja. La segunda razón es más práctica y se trata de integrales reales que pueden evaluarse mediante integración compleja, en tanto que los métodos usuales del Cálculo real no son útiles para esto.

Siendo una generalización natural de la integral real, la integral compleja a lo largo de una curva se define de la siguiente manera.

Sea $C$ una **curva seccionalmente suave**, que consta de un número finito de curvas suaves en el plano complejo $z$ y representada en forma paramétrica:

$$z(t) = x(t) + iy(t), \quad a \leq t \leq b \tag{3.55}$$

Por convención, la dirección positiva a lo largo de $C$ se escoge al sentido de los valores crecientes del parámetro $t$. Sea $f(z)$ una función compleja continua en $C$. Partiendo el intervalo $a \leq t \leq b$ en $n$ subintervalos en los puntos $t_0 (= a) < t_1 < t_2 < \cdots < t_n (= b)$, se obtiene la partición de la curva $C$ mediante los puntos $z_0, z_1, z_2, \ldots, z_n$, en donde $z_k = z(t_k)$. La **integral de línea** de $f(z)$ a lo largo de una curva orientada $C$ se define como sigue:

$$\int_C f(z)\, dz \equiv \lim_{n\to\infty, \max \Delta z_k \to 0} \sum_{k=0}^{n} f(\zeta_k) \Delta z_k \tag{3.56}$$

en donde $\Delta z_k = z_k - z_{k-1}$ y $\zeta_k$ es un punto en el segmento $k$-ésimo de $C$.

Sean $f(z) = u(x, y) + iv(x, y)$, $\zeta_k = \xi_k + i\eta_k$, $\Delta z_k = \Delta x_k + i\Delta y_k$. Entonces, de la ecuación (3.56) se sigue que la integral compleja de línea se representa en términos de integrales reales:

$$\int_C f(z)\, dz = \int_C (udx - vdy) + i \int_C (udy + vdx) \tag{3.57}$$

Esto demuestra que la integral de línea, ecuación (3.56) existe y su valor es independiente de la partición de la curva $C$ y los puntos intermedios $\zeta_k$. La curva $C$ recibe el nombre de **trayectoria de integración**.

Al aplicar la representación paramétrica, ecuación (3.55), la integral de línea toma la forma,

$$\begin{aligned} \int_C f(z)\,dz &= \int_C (udx - vdy) + i\int_C (udy + vdx) \\ &= \int_a^b (u\dot{x} - v\dot{y})\,dt + i\int_a^b (u\dot{y} + v\dot{x})\,dt \end{aligned} \tag{3.58}$$

donde $u = u(x(t), y(t))$, $v = v(x(t), y(t))$ y el punto denota la derivada con respecto al parámetro $t$.

**Ejemplo 3.15.** *Con el fin de mostrar la aplicación de la fórmula (3.57), evaluamos la integral*

$$\int_C (z^*)^2\,dz$$

*a lo largo de la línea recta que conecta los puntos $z = 0$ y $z = 1 + 2i$. Dado que $(z^*)^2 = x^2 - y^2 - i2xy$, al sustituir $u = x^2 - y^2$ y $v = -2xy$ en la ecuación (3.57), se tiene*

$$\int_C (z^*)^2\,dz = \int_C \left(\left(x^2 - y^2\right)dx + 2xydy\right) + i\int_C \left(\left(x^2 - y^2\right)dy - 2xydx\right)$$

*Al usar la representación cartesiana de $C$, $y = 2x$, la fórmula anterior se reduce a la evaluación de dos integrales definidas:*

$$\int_C (z^*)^2\,dz = \int_0^1 5x^2dx + i\int_0^1 \left(-10x^2\right)dx = \frac{5}{3} - i\frac{10}{3}$$

**Ejemplo 3.16.** *La integral*

$$\oint_C \frac{dz}{z}$$

*en donde $C$ es el círculo unitario con centro en $z = 0$, es más fácil de evaluar usando la representación paramétrica, ecuación (3.58). La representación paramétrica de $C$ es*

$$z(t) = \cos t + i\operatorname{sen} t = e^{it}, \quad 0 \le t \le 2\pi$$

*Por tanto, $dz = ie^{it}dt$ y, a partir de esto, se obtiene la integral*

$$\oint_C \frac{dz}{z} = \int_0^{2\pi} \frac{ie^{it}dt}{e^{it}} = 2\pi i$$

Basándose en la definición de integral compleja, ecuación (3.56), es fácil obtener las propiedades siguientes.

Si la trayectoria de integración $C$ está compuesta de dos partes, $C_1$ y $C_2$, entonces

$$\int_{C_1 \cup C_2} f(z)\, dz = \int_{C_1} f(z)\, dz + \int_{C_2} f(z)\, dz \tag{3.59}$$

Si se invierte el sentido de la integración, se cambia el signo de la integral:

$$\int_{z_0}^{z_1} f(z)\, dz = -\int_{z_1}^{z_0} f(z)\, dz \tag{3.60}$$

La integral es lineal con respecto al integrando, esto es,

$$\int_C (k_1 f_1(z) + k_2 f_2(z))\, dz = k_1 \int_C f_1(z)\, dz + k_2 \int_C f_2(z)\, dz \tag{3.61}$$

Para enunciar el teorema de Cauchy, que es de gran importancia en el análisis complejo, se requieren los conceptos siguientes. Una **curva simple** es continua y no tiene autointersecciones. Una **curva simple cerrada** tiene una sola autointersección. Se dice que un dominio $D$ en el plano complejo es **simplemente conexo** (**de conectividad simple**) si toda curva simple cerrada en $D$ sólo encierra puntos de éste, es decir, cualquier curva simple cerrada en $D$ puede reducirse por una transformación continua a un punto sin cruzar la frontera del dominio. Es claro que el interior de una curva simple cerrada es un dominio simplemente conexo. Un dominio que no es simplemente conexo es **de conectividad múltiple**. Un anillo circular es el ejemplo de dominio doblemente conexo. Se dice que un dominio $D$ es **acotado** si se encuentra dentro de un círculo de radio finito trazado alrededor del origen.

En el análisis real se demuestra que la igualdad

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x} \tag{3.62}$$

es la condición suficiente y necesaria para que la integral del tipo

$$\int_C (M(x,y)\, dx + N(x,y)\, dy) \tag{3.63}$$

sea independiente de la trayectoria de integración en un dominio $D$ cerrado y simplemente conexo a condición de que $M(x,y)\, dx$ y $N(x,y)$ y las primeras derivadas parciales en la ecuación (3.62) sean funciones continuas en $D$. En realidad, dicha condición asegura que la forma diferencial $Mdx + Ndy$ es una diferencial exacta y esto es que la integral es independiente de la trayectoria de integración. Sea $f(z) = u(x,y) + iv(x,y)$ una función analítica en $D$, esto es, $u(x,y)$ y $v(x,y)$ satisfacen las ecuaciones de Cauchy-Riemann. Entonces, la parte real e imaginaria de la integral (3.57) cumplen con la condición (3.62) y, por tanto, la integral (3.57) de una función analítica es independiente de la trayectoria de integración. Si la trayectoria de integración es una curva simple cerrada en un dominio simplemente conexo $D$, entonces el valor de la integral es cero. Este resultado fundamental se conoce como el teorema de Cauchy.

**Teorema 3.3.** *(Teorema integral de Cauchy) Sea $f(z)$ una función analítica en un dominio $D$ simplemente conexo y acotado. Entonces,*

$$\oint_C f(z)\,dz = 0 \tag{3.64}$$

*para toda trayectoria $C$ simple cerrada en $D$.*

Cabe mencionar que además de la demostración anterior, existen otras demostraciones del teorema de Cauchy; una de éstas es la demostración original de Cauchy que usa el teorema de Green y las ecuaciones de Cauchy-Riemann.

Una consecuencia de gran interés práctico del teorema integral de Cauchy es el **principio de deformación de la trayectoria de integración**. Suponemos que la curva de integración simple cerrada $C$ en la ecuación (3.64) se partirá en dos arcos $C_1$ y $\tilde{C}_2$ de manera que $C = C_1 \cup \tilde{C}_2$ y el arco $C_1$ une dos puntos de partición, digamos $z_1$ y $z_2$, con el sentido de recorrido de $z_1$ a $z_2$; el arco $\tilde{C}_2$ une $z_2$ y $z_1$, con el sentido de recorrido de $z_2$ a $z_1$. De las ecuaciones (3.59) y (3.64), se tiene que

$$\oint_C f(z)\,dz = \int_{C_1(z_1\to z_2)} f(z)\,dz + \int_{\tilde{C}_2(z_2\to z_1)} f(z)\,dz = 0$$

El cambio del sentido de la integración conduce al cambio del signo de la integral. Por tanto, de la ecuación anterior se sigue que

$$\int_{C_1(z_1\to z_2)} f(z)\,dz = -\int_{\tilde{C}_2(z_2\to z_1)} f(z)\,dz = \int_{C_2(z_1\to z_2)} f(z)\,dz \tag{3.65}$$

en donde $C_2$ es la curva que coincide con $\tilde{C}_2$, pero tiene la orientación opuesta. Entonces, si $f(z)$ es analítica en un dominio $D$ simplemente conexo, la integral de línea entre dos puntos no depende de la trayectoria que une estos puntos, siempre y cuando las trayectorias de integración pertenezcan a $D$. Nótese que, en realidad, este principio de deformación de la trayectoria lo hemos usado para demostrar el teorema de Cauchy-Riemann.

### 3.3.3. Teorema integral de Cauchy para dominios de conectividad múltiple

La suposición fundamental para establecer el teorema de Cauchy es que el dominio de analiticidad de la función a integrar es simplemente conexo. La extensión del teorema a los dominios de conectividad múltiple es como sigue.

Nótese que se puede cortar un dominio de conectividad múltiple de modo que el dominio resultante se vuelve simplemente conexo. Para ser más específico, considérese un dominio cerrado $D$ doblemente conexo que está limitado por las curvas simples cerradas $C$ de su exterior y $C_1$ de su interior, de modo que la última se encuentra completamente por dentro de $C$, como se muestra en la figura 3.4. Se supone que una función $f(z)$ es analítica en $D$. Un corte $AB$ que une un punto $A$ en $C$ con un punto $B$ en $C_1$ vuelve el dominio $D$ en un dominio

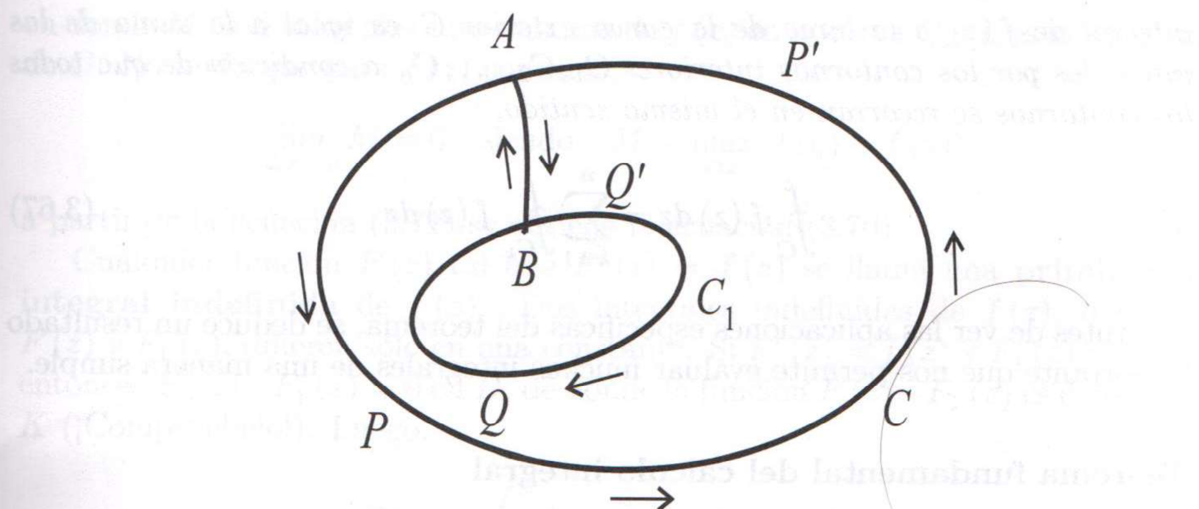



Figura 3.4. Dominio doblemente conexo

simplemente conexo con su frontera formada por las curvas $C$, $AB$, $C_1$ y $BA$. A este dominio se aplica el teorema de Cauchy,

$$\oint_{C(APP'A)} f(z)\,dz + \int_{AB} f(z)\,dz + \oint_{C_1(BQ'QB)} f(z)\,dz + \int_{BA} f(z)\,dz = 0$$

en donde los subíndices indican las direcciones de integración a lo largo de $C$, $AB$, $C_1$ y $BA$. Las integrales por el corte en dos direcciones opuestas, $AB$ y $BA$, se cancelan. Cambiando la dirección de integración a lo largo del contorno $C_1$ de la fórmula anterior, se tiene

$$\oint_{C(APP'A)} f(z)\,dz = \oint_{C_1(BQQ'B)} f(z)\,dz \tag{3.66}$$

Se ve que los valores de las integrales de una función analítica $f(z)$ a lo largo de dos diferentes contornos, $C$ y $C_1$, con el mismo sentido de recorrido, son iguales entre sí, pero no son necesariamente iguales a cero, puesto que $f(z)$ podría ser no analítica en los puntos del dominio encerrado por $C_1$. De esta observación de inmediato se sigue el importante **principio de deformación de contornos**: la integral de una función analítica $f(z)$ a lo largo de una curva simple cerrada (un **contorno**) $C$ tiene el mismo valor de la integral a lo largo de cualquier otro contorno $C_1$ que se obtiene de $C$ mediante una deformación continua sin pasar por los puntos singulares de $f(z)$.

En el caso de un dominio más complicado se requieren varios cortes, que no son únicos, para convertirlo en un dominio simplemente conexo. Luego, el resultado anterior se generaliza en el siguiente teorema.

**Teorema 3.4.** ***(Teorema integral de Cauchy para dominios múltiplemente conexos)*** *Sea $f(z)$ una función analítica en un dominio cerrado múltiplemente conexo $D$ que está limitado de su exterior por una curva simple cerrada $C$ y por unas curvas simples cerradas interiores $C_1, C_2, \ldots, C_n$. Entonces, la*

*integral de $f(z)$ a lo largo de la curva exterior $C$ es igual a la suma de las integrales por los contornos interiores $C_1, C_2, \ldots, C_n$ a condición de que todos los contornos se recorran en el mismo sentido,*

$$\oint_C f(z)\,dz = \sum_{k=1}^{n} \oint_{C_k} f(z)\,dz \tag{3.67}$$

Antes de ver las aplicaciones específicas del teorema, se deduce un resultado importante que nos permite evaluar muchas integrales de una manera simple.

**Teorema fundamental del cálculo integral**

Sean $f(z)$ una función analítica en un dominio simplemente conexo $D$ y $C$ una curva que conecta dos puntos $z_0$ y $z$ de este dominio. Considérese la integral

$$F(z) = \int_{z_0}^{z} f(\zeta)\,d\zeta \tag{3.68}$$

a lo largo de $C$, siendo $z_0$ cualquier punto fijo en $D$. Puesto que $f(z)$ es una función analítica en $D$, la integral no depende de la trayectoria de integración y está completamente determinada por los puntos $z_0$ y $z$. Por tanto, la integral (3.68) define una función $F(z)$. Se comprueba que ésta es una función analítica de la variable $z$ en $D$ y que

$$F'(z) = f(z) \tag{3.69}$$

evaluamos la razón,

$$\begin{aligned}
\frac{F(z+\Delta z) - F(z)}{\Delta z} &= \frac{1}{\Delta z}\left[\int_{z_0}^{z+\Delta z} f(\zeta)\,d\zeta - \int_{z_0}^{z} f(\zeta)\,d\zeta\right] \\
&= \frac{1}{\Delta z}\int_{z}^{z+\Delta z} f(\zeta)\,d\zeta \\
&= \frac{1}{\Delta z}\int_{z}^{z+\Delta z} \left(f(\zeta) - f(z) + f(z)\right)d\zeta \\
&= f(z) + \frac{1}{\Delta z}\int_{z}^{z+\Delta z} \left(f(\zeta) - f(z)\right)d\zeta
\end{aligned}$$

Resulta que

$$\lim_{\Delta z \to 0} \frac{1}{\Delta z}\int_{z}^{z+\Delta z} \left(f(\zeta) - f(z)\right)d\zeta = 0 \tag{3.70}$$

y, por consiguiente, es válida la ecuación (3.69). Para evaluar el límite anterior se usa la siguiente estimación del valor absoluto de la integral de línea:

$$\left|\int_C f(z)\,dz\right| \leq \int_C |f(z)|\,|dz| \leq ML \tag{3.71}$$

en donde $M$ es el valor absoluto máximo de $f(z)$ en la curva $C$ y $L$ es la longitud de $C$. Notando que para $f(z)$ continua,

$$\lim_{\Delta z \to 0} M = 0 \quad \text{siendo} \quad M = \max_{\Delta z} |f(\zeta) - f(z)|$$

a partir de la ecuación (3.71) se obtiene la ecuación (3.70).

Cualquier función $F(z)$ tal que $F'(z) = f(z)$ se llama una **primitiva** o **integral indefinida** de $f(z)$. Dos integrales indefinidas de $f(z)$, digamos $F(z)$ y $F_1(z)$, difieren sólo en una constante. Si $F'(z) = f(z)$ y $F_1'(z) = f(z)$, entonces $F'(z) - F_1'(z) = 0$ en $D$, de donde la función $F(z) - F_1(z)$ es constante $K$ (¡Compruébelo!). Luego,

$$F(z) = \int_{z_0}^{z} f(\zeta)\, d\zeta = F_1(z) + K$$

de donde $F(z_0) = F_1(z_0) + K = 0$ y, por consiguiente,

$$F(z) = \int_{z_0}^{z} f(\zeta)\, d\zeta = F_1(z) - F_1(z_0) \tag{3.72}$$

Este resultado se conoce como el **teorema fundamental del cálculo integral**.

**Teorema 3.5.** *Si $f(z)$ es analítica en un dominio simplemente conexo $D$, entonces existe una integral indefinida de $f(z)$ en $D$, es decir, una función analítica $F(z)$ tal que $F'(z) = f(z)$. Además, para cualquier trayectoria $C$ que une dos puntos $z_0$ y $z$ en $D$, el valor de la integral de línea de $f(z)$ es igual a la diferencia de valores de la integral indefinida en los puntos extremos de la trayectoria de integración, ecuación (3.72).*

**Ejemplo 3.17.** *Sea $C$ cualquier trayectoria que une dos puntos $z = 0$ y $z = 2 + i$. Entonces,*

$$\int_C \left(3z^2\right) dz = 3 \int_0^{2+i} z^2 dz = z^3 \Big|_0^{2+i} = (2+i)^3$$

**Ejemplo 3.18.** *Sea $C$ cualquier trayectoria que une dos puntos $z = i$ y $z = \pi/2$. De manera similar,*

$$\int_C \cos z dz = \operatorname{sen} z \Big|_i^{\pi/2} = \operatorname{sen} \frac{\pi}{2} - \operatorname{sen} i = 1 - i \operatorname{senh} 1$$

### 3.3.4. Fórmula integral de Cauchy

La consecuencia importante del teorema integral de Cauchy es la fórmula integral de Cauchy, que establece el hecho de que toda función $f(z)$ analítica en un dominio cerrado simplemente conexo $\bar{D} = D \cup C$ está completamente determinada en $D$ por valores de $f(z)$ en la frontera $C$ del dominio $D$.

**Corolario.** *Sea $f(z)$ una función analítica en un dominio cerrado simplemente conexo $\bar{D} = D \cup C$ y la frontera $C$ sea una curva simple cerrada.*

*Entonces, para cualquier punto $z \in D$ es válida la* **fórmula integral** *(representación integral)* **de Cauchy**

$$f(z) = \frac{1}{2\pi i} \oint_C \frac{f(\zeta)\, d\zeta}{\zeta - z} \tag{3.73}$$

*tomándose la integración en sentido contrario de las manecillas del reloj.*

**Demostración.** *La función*

$$\frac{f(z)}{z - z_0}$$

*es analítica en cada punto interior de $D$ excepto, posiblemente, el punto $z = z_0$. Encerrando el punto $z_0$ por un pequeño círculo $\gamma$, de radio $\rho$ y con el centro en $z_0$, del principio de deformación de contornos, ecuación (3.66), se tiene que*

$$\begin{aligned} \oint_C \frac{f(z)\, dz}{z - z_0} &= \oint_\gamma \frac{f(z)\, dz}{z - z_0} \\ &= f(z_0) \oint_\gamma \frac{dz}{z - z_0} + \oint_\gamma \frac{f(z) - f(z_0)}{z - z_0} dz \end{aligned} \tag{3.74}$$

*en donde se usó la transformación idéntica $f(z) = f(z_0) - f(z_0) + f(z)$.*

*Con el cambio de la variable de integración en el círculo $\gamma$ de la siguiente manera, $\zeta = z - z_0 = \rho e^{i\varphi}$ y $d\zeta = dz = \rho i e^{i\varphi} d\varphi$, se tiene*

$$\oint_\gamma \frac{dz}{z - z_0} = \oint_\gamma \frac{d\zeta}{\zeta} = \int_0^{2\pi} \frac{\rho i e^{i\varphi} d\varphi}{\rho e^{i\varphi}} = 2\pi i \tag{3.75}$$

*Sea*

$$M = \max_\gamma |f(z) - f(z_0)|$$

*Recordando que $|z - z_0| = \rho$ y $|dz| = \rho d\varphi$ en el círculo $\gamma$, se obtiene la siguiente estimación:*

$$\left| \oint_\gamma \frac{f(z) - f(z_0)}{z - z_0} dz \right| \leq \oint_\gamma |f(z) - f(z_0)| \left| \frac{dz}{z - z_0} \right| \leq M \int_0^{2\pi} d\varphi = 2\pi M \tag{3.76}$$

*Además,*

$$\lim_{\rho \to 0} M = 0$$

*siendo $f(z)$ analítica y, por lo tanto, continua. Sustituyendo las ecuaciones (3.75) y (3.76) en la ecuación (3.74) y tendiendo $\rho \to 0$, se obtiene la representación integral de Cauchy*

$$\oint_C \frac{f(z)\, dz}{z - z_0} = 2\pi i f(z_0)$$

*para la función $f(z)$ en cualquier punto $z_0 \in D$. Esta fórmula nos permite calcular el valor de $f(z)$ en cualquier punto interno $z_0 \in D$ a través de los*

*valores de $f(z)$ en el contorno de frontera $C$. Denominando la variable de integración por $\zeta$ y el punto interno en $D$ por $z$, de la fórmula anterior se obtiene la ecuación (3.73).* ■

En el caso de un dominio de conectividad múltiple, la fórmula integral de Cauchy se extiende de manera natural, haciendo los cortes que convierten este dominio en uno simplemente conexo. El resultado final es como sigue:

$$f(z) = \frac{1}{2\pi i}\left[\oint_C \frac{f(\zeta)\,d\zeta}{\zeta - z} - \sum_{k=1}^{n}\oint_{C_k}\frac{f(\zeta)\,d\zeta}{\zeta - z}\right] \tag{3.77}$$

en donde las integrales se toman en sentido contrario al de las manecillas del reloj y $z_0$ es cualquier punto interno de un dominio múltiplemente conexo $D$, limitado por el contorno externo $C$ y por los contornos internos $C_1, C_2, ..., C_n$.

Basándose en la fórmula integral de Cauchy, ecuación (3.73), se puede comprobar que una función analítica tiene no sólo la primera derivada sino las derivadas de todos los órdenes. Sea $f(z)$ una función analítica en un dominio simplemente conexo $D$. Entonces es válida la ecuación (3.73) y existe la primera derivada de $f(z)$, es decir,

$$f'(z) = \lim_{\Delta z \to 0}\frac{f(z+\Delta z) - f(z)}{\Delta z}$$

Por otro lado, de la ecuación (3.73) se tiene

$$\begin{aligned}\lim_{\Delta z\to 0}\frac{f(z+\Delta z)-f(z)}{\Delta z} &= \lim_{\Delta z\to 0}\frac{1}{2\pi i\Delta z}\left[\oint_C\frac{f(\zeta)\,d\zeta}{\zeta-(z+\Delta z)} - \oint_C\frac{f(\zeta)\,d\zeta}{\zeta - z}\right]\\ &= \lim_{\Delta z\to 0}\frac{1}{2\pi i}\left[\oint_C\frac{f(\zeta)\,d\zeta}{(\zeta - z-\Delta z)(\zeta - z)}\right]\end{aligned}$$

Tomando el límite bajo el signo de integral, que es legítimo si $f(z)$ es continua, se obtiene la primera derivada

$$f'(z) = \frac{1}{2\pi i}\oint_C\frac{f(\zeta)\,d\zeta}{(\zeta - z)^2}$$

De manera similar, se comprueba la existencia de las derivadas de todos los órdenes,

$$f^{(n)}(z) = \frac{n!}{2\pi i}\oint_C\frac{f(\zeta)\,d\zeta}{(\zeta - z)^{n+1}}, \quad n = 0, 1, 2, 3, \ldots \tag{3.78}$$

Nótese que la ecuación (3.78) se obtiene al derivar formalmente la fórmula de Cauchy, ecuación (3.73), con respecto a $z$.

**Ejemplo 3.19.** *Notando que el punto $z = -1$ se encuentra dentro del contorno $|z| = 2$ y que la función $e^{-z}$ es analítica, se tiene que*

$$\oint_C\frac{e^{-z}dz}{z+1} = 2\pi i e^{-z}\mid_{z=-1} = e2\pi i$$

**Ejemplo 3.20.** *Para el contorno $|z| = 2$, de la ecuación (3.78) se tiene*

$$\oint_C \frac{z^2 dz}{(z-i)^2} = 2\pi i\,(2z)\,|_{z=i} = -4\pi$$

### 3.3.5. Series de Taylor

En esta sección consideraremos la representación de funciones analíticas en series de potencias, que es de gran importancia en el análisis complejo. Recordaremos algunos de los conceptos básicos de la convergencia de series. Si una serie de potencias $\sum_{k=0}^{\infty} a_k z^k$ converge para $z = z_1$, ésta converge absoluta y uniformemente en cualquier disco circular $|z| \leq r$, donde $r < |z_1|$. El círculo de radio $r$, tal que la serie $\sum_{k=0}^{\infty} a_k z^k$ converge para toda $z$ del disco $|z| < r$ y diverge para algunos puntos $|z| > r$, se llama **círculo de convergencia** y su radio $r$ es el **radio de convergencia** de la serie. El radio de convergencia a menudo se determina mediante la prueba de la razón,

$$r = \lim_{k\to\infty} \left| \frac{a_{k-1}}{a_k} \right| \tag{3.79}$$

siempre y cuando este límite exista.

**Ejemplo 3.21.** *Para la serie*

$$\sum_{k=1}^{\infty} \frac{(-1)^k z^k}{k}$$

*el radio de convergencia es $r = 1$, porque*

$$r = \lim_{k\to\infty} \left| \frac{a_{k-1}}{a_k} \right| = \lim_{k\to\infty} \frac{k}{k-1} = 1$$

**Ejemplo 3.22.** *La serie*

$$\sum_{k=0}^{\infty} \frac{z^k}{k!}$$

*converge en todo el plano complejo $z$, porque su radio de convergencia es infinito,*

$$r = \lim_{k\to\infty} \left| \frac{a_{k-1}}{a_k} \right| = \lim_{k\to\infty} \frac{k!}{(k-1)!} = \infty$$

Suponiendo que $f(z)$ es una función analítica en $z = a$, ésta tiene derivadas de todos los órdenes $f^{(k)}(a)$, $k = 0, 1, 2, \ldots$ Por tanto, podemos asociar con $f(z)$ la serie de potencias

$$S(z) \equiv \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (z-a)^k$$

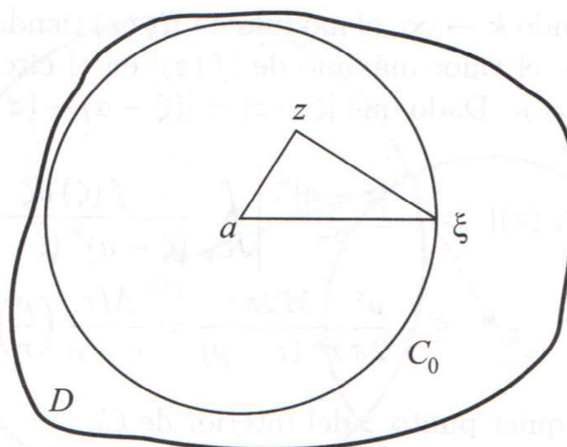



Figura 3.5. Dominio para el desarrollo en serie de Taylor

que converge en un disco circular $|z-a|<r$. La pregunta es: ¿esta serie converge a la función $f(z)$? Resulta que las funciones analíticas siempre pueden representarse por series de potencias.

Sea $f(z)$ analítica en un dominio $D$ y sea $C_0$ un círculo, de radio $r$ y con el centro en $z=a$, que se encuentra completamente en $D$. Si $z$ es un punto interior del disco $C_0$ (véase la figura 3.5), entonces, de la fórmula integral de Cauchy, se tiene

$$\begin{aligned} f(z) &= \frac{1}{2\pi i}\oint_{C_0}\frac{f(\zeta)\,d\zeta}{\zeta - z} \\ &= \frac{1}{2\pi i}\oint_{C_0}\frac{f(\zeta)\,d\zeta}{(\zeta-a)\left[1-(z-a)/(\zeta-a)\right]} \end{aligned} \tag{3.80}$$

A partir de la progresión geométrica (o por división) se obtiene la relación

$$\frac{1}{1-t} = 1+t+t^2+\cdots+t^{k-1}+\frac{t^k}{1-t}$$

Sustituyendo esta expresión con $t=(z-a)/(\zeta-a)$ en la ecuación (3.80) se obtiene que

$$f(z)=\frac{1}{2\pi i}\left[\begin{array}{c}\oint_{C_0}\frac{f(\zeta)d\zeta}{\zeta-a}+(z-a)\oint_{C_0}\frac{f(\zeta)d\zeta}{(\zeta-a)^2}+\cdots \\ +(z-a)^{k-1}\oint_{C_0}\frac{f(\zeta)d\zeta}{(\zeta-a)^k}\end{array}\right]+R_k(z)$$

donde

$$R_k(z)=\frac{(z-a)^k}{2\pi i}\oint_{C_0}\frac{f(\zeta)\,d\zeta}{(\zeta-a)^k(\zeta-z)}$$

Utilizando la ecuación (3.78), este desarrollo puede escribirse en la forma

$$\begin{aligned} f(z) &= f(a)+f'(a)(z-a)+\frac{f''(a)}{2!}(z-a)^2+\cdots \\ &\quad +\frac{f^{(k-1)}(a)}{(k-1)!}(z-a)^{k-1}+R_k(z) \end{aligned} \tag{3.81}$$

En el límite, cuando $k \to \infty$, el módulo de $R_k(z)$ tiende a cero. Para mostrar esto, suponemos que el valor máximo de $|f(z)|$ en el círculo $C_0$ es $M$, el radio de $C_0$ es $r$ y $|z-a| = \rho$. Dado que $|\zeta - z| = |(\zeta - a) - (z-a)| \geq r - \rho$, se tiene

$$\begin{aligned} |R_k(z)| &= \frac{|z-a|^k}{2\pi} \left| \oint_{C_0} \frac{f(\zeta)\, d\zeta}{(\zeta - a)^k (\zeta - z)} \right| \\ &\leq \frac{\rho^k}{2\pi} \frac{M 2\pi r}{r^k (r-\rho)} = \frac{Mr}{r-\rho} \left(\frac{\rho}{r}\right)^k \end{aligned}$$

Entonces, para cualquier punto $z$ del interior de $C_0$

$$\lim_{k\to\infty} |R_k(z)| = 0$$

dado que $(\rho/r) < 1$. En consecuencia, de la ecuación (3.81) se obtiene la serie de potencias

$$f(z) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (z-a)^k \tag{3.82}$$

que converge a la función $f(z)$ en cada punto interior de $|z-a| = r$. Esta serie se llama **serie de Taylor** de la función analítica $f(z)$ expandida en el punto $z = a$. En el caso particular $a = 0$, esta serie se conoce como **serie de Maclaurin** de $f(z)$. Se puede comprobar que la representación (3.82) es única.

El radio de convergencia $r$ de la serie (3.82) es igual a la distancia desde el punto $z = a$ hasta el punto singular más cercano, $z = z_0$, de $f(z)$, es decir, $r = |z_0 - a|$. Por supuesto, la función $f(z)$ puede ser analítica en un punto $z = a_1$ fuera del círculo $C_0$. El dominio de analiticidad $D$ de la función $f(z)$ puede ser completamente cubierto por un conjunto de círculos $\{C_0, C_1, C_2, \ldots\}$ cuyos centros $a_0, a_1, a_2, \ldots$ son puntos internos de $D$ y el radio $r_i$ de cada círculo es la distancia desde su centro $a_i$ hasta el punto singular más cercano, $z = z_i$, de $f(z)$. En cada círculo $C_i$ la función $f(z)$ puede representarse por su serie de Taylor

$$f(z) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a_i)}{k!} (z-a_i)^k, \quad i = 0, 1, 2, \ldots \tag{3.83}$$

Si dos círculos se traslapan entre sí, entonces las dos series correspondientes convergen a $f(z)$ en el dominio de traslape, es decir, la función $f(z)$ puede representarse mediante diferentes series de Taylor en el dominio de traslape de diferentes círculos $C_i$.

La situación escrita está estrechamente relacionada con la extensión de una función analítica definida en un dominio a otro dominio más grande. Sea $f_1(z)$ una función analítica definida por medio de su serie de Taylor en un círculo $C_1$(véase la figura 3.6),

$$f_1(z) = \sum_{k=0}^{\infty} \frac{f_1^{(k)}(a_1)}{k!} (z-a_1)^k$$

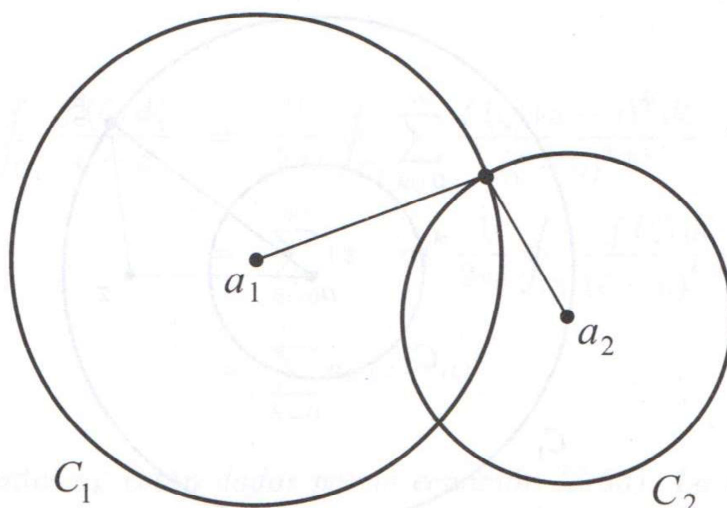



Figura 3.6. Continuación analítica mediante series de Taylor

Si la función $f_1(z)$ es analítica en un punto $z = a_2$ fuera de $C_1$ y la serie

$$S_2(z) = \sum_{k=0}^{\infty} \frac{f_1^{(k)}(a_2)}{k!}(z - a_2)^k$$

es convergente en un círculo $C_2$, ésta define una función analítica $f_2(z) = S_2(z)$ en el interior de $C_2$. Si, además, los círculos $C_1$ y $C_2$ se traslapan entre sí, entonces $f_1(z) = f_2(z)$ en el dominio de traslape y se dice que $f_2(z)$ es una **continuación analítica** de $f_1(z)$ desde el dominio limitado por $C_1$ al dominio limitado por $C_2$. Estas continuaciones son de gran importancia para las aplicaciones de la teoría de funciones analíticas en la solución de problemas de eigenvalores.

**Ejercicio.** *Expandir la función* $f(z) = 1/(1 - z)$ *en series de Taylor alrededor de los puntos* $a_1 = 0$, $a_2 = -1$, $a_3 = i$. *Dibujar círculos de convergencia para cada serie.*

### 3.3.6. Expansión de Laurent

En la sección anterior demostramos que una función $f(z)$ analítica en un punto dado $a$ puede ser representada en la vecindad de éste mediante una serie de potencias, la serie de Taylor. Además, esta serie representa $f(z)$ en un dominio circular de radio que es igual a la distancia de $a$ hasta el punto singular más cercano de la función $f(z)$. En varias aplicaciones es necesario desarrollar una función $f(z)$ alrededor de los puntos en donde es singular. En tales casos no existe la serie de Taylor y se necesita la **serie de Laurent**. La serie de Laurent representa una función $f(z)$ que es analítica en un anillo circular limitado por dos círculos concéntricos, $C_1$ exterior y $C_2$ interior (véase la figura 3.7). Como en el caso de la serie de Taylor, $f(z)$ puede ser singular en algunos puntos fuera de $C_1$ y, como característica nueva, puede ser singular en algunos puntos dentro de $C_2$.

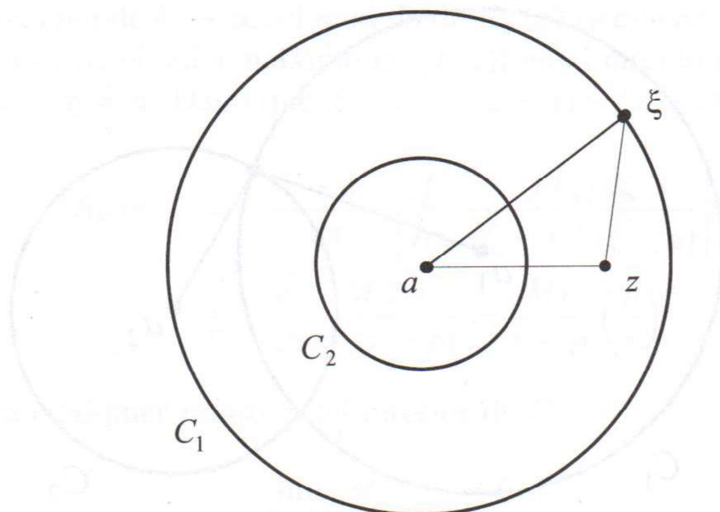



Figura 3.7. Dominio para el desarrollo en serie de Laurent

**Teorema 3.6.** *Una función $f(z)$ analítica en un anillo circular $R_2 \leq |z-a| \leq R_1$ puede ser representada en cada punto interior de éste por medio de la serie de Laurent*

$$f(z) = \sum_{k=0}^{\infty} a_k (z-a)^k + \sum_{k=1}^{\infty} \frac{a_{-k}}{(z-a)^k} \tag{3.84}$$

*donde*

$$a_k = \frac{1}{2\pi i} \oint_{C_1} \frac{f(\zeta)\, d\zeta}{(\zeta-a)^{k+1}}, \quad k = 0, 1, 2, \ldots \tag{3.85}$$

$$a_{-k} = \frac{1}{2\pi i} \oint_{C_2} \frac{f(\zeta)\, d\zeta}{(\zeta-a)^{-k+1}}, \quad k = 1, 2, \ldots \tag{3.86}$$

*siendo $C_1$ y $C_2$ fronteras del anillo circular y tomándose cada una de las integrales en sentido contrario al de las manecillas del reloj.*

**Demostración.** *Según la fórmula integral de Cauchy, ecuación (3.77), para cualquier punto $z$ del interior del anillo $R_2 \leq |z-a| \leq R_1$ se tiene*

$$f(z) = \frac{1}{2\pi i} \oint_{C_1} \frac{f(\zeta)\, d\zeta}{\zeta - z} - \frac{1}{2\pi i} \oint_{C_2} \frac{f(\zeta)\, d\zeta}{\zeta - z} \tag{3.87}$$

*donde se integra en sentido contrario al de las manecillas del reloj.*

*Evaluamos la integral por la circunferencia $C_1$. Para cualquier $\zeta$ de $C_1$ se cumple $|z-a| / |\zeta - a| < 1$ y, por tanto, usando la fórmula de progresión geométrica se obtiene*

$$\frac{1}{\zeta - z} = \frac{1}{\zeta - a} \frac{1}{1-(z-a)/(\zeta-a)} = \frac{1}{\zeta - a} \sum_{k=0}^{\infty} \left( \frac{z-a}{\zeta-a} \right)^k$$

*Luego,*

$$\begin{aligned}\frac{1}{2\pi i}\oint_{C_1}\frac{f(\zeta)\,d\zeta}{\zeta-z} &= \frac{1}{2\pi i}\oint_{C_1}\sum_{k=0}^{\infty}\frac{f(\zeta)(z-a)^k\,d\zeta}{(\zeta-a)^{k+1}} \\ &= \sum_{k=0}^{\infty}(z-a)^k\frac{1}{2\pi i}\oint_{C_1}\frac{f(\zeta)\,d\zeta}{(\zeta-a)^{k+1}} \\ &= \sum_{k=0}^{\infty}a_k(z-a)^k\end{aligned}$$

*donde los coeficientes $a_k$ están dados por la ecuación (3.85). La integración de la serie término por término se justifica idénticamente al caso de las series de Taylor.*

*Consideremos la integral por la circunferencia $C_2$. En este caso se tiene $|\zeta-a|/|z-a|<1$ y, por tanto,*

$$\frac{1}{\zeta-z}\equiv-\frac{1}{z-a}\frac{1}{1-(\zeta-a)/(z-a)}=-\frac{1}{z-a}\sum_{k=0}^{\infty}\left(\frac{\zeta-a}{z-a}\right)^k$$

*Entonces,*

$$\begin{aligned}-\frac{1}{2\pi i}\oint_{C_2}\frac{f(\zeta)\,d\zeta}{\zeta-z} &= \frac{1}{2\pi i}\oint_{C_2}\sum_{k=0}^{\infty}\frac{f(\zeta)(\zeta-a)^k\,d\zeta}{(z-a)^{k+1}} \\ &= \frac{1}{2\pi i}\oint_{C_2}\sum_{k=1}^{\infty}\frac{f(\zeta)(\zeta-a)^{k-1}\,d\zeta}{(z-a)^k} \\ &= \sum_{k=1}^{\infty}\frac{1}{(z-a)^k}\frac{1}{2\pi i}\oint_{C_2}f(\zeta)(\zeta-a)^{k-1}\,d\zeta \\ &= \sum_{k=1}^{\infty}\frac{a_{-k}}{(z-a)^k}\end{aligned}$$

*donde los coeficientes $a_{-k}$ se determinan por la ecuación (3.86). La integración de la serie término por término se justifica de manera similar al caso de las series de Taylor.*

*Queda completamente demostrado el teorema.* ■

Cuando $f(z)$ es analítica en el interior del círculo $C_2$, el integrando en la ecuación (3.86) es una función analítica también. Entonces, $a_{-k}=0$ según el teorema de Cauchy. En este caso

$$a_k=\frac{1}{2\pi i}\oint_{C_1}\frac{f(\zeta)\,d\zeta}{(\zeta-a)^{k+1}}=\frac{f^{(k)}(a)}{k!},\quad k=0,1,2,\ldots$$

y, por tanto, la serie de Laurent se reduce a la serie de Taylor.

Nótese que la serie de Laurent, ecuación (3.84), puede escribirse en forma más compacta,

$$f(z) = \sum_{k=-\infty}^{\infty} a_k (z-a)^k \tag{3.88}$$

donde los coeficientes $a_k$ se calculan por la fórmula

$$a_k = \frac{1}{2\pi i} \oint_C \frac{f(\zeta)\, d\zeta}{(\zeta - a)^{k+1}}, \quad k = 0, \pm 1, \pm 2, \ldots \tag{3.89}$$

La ecuación (3.89) se obtiene de las ecuaciones (3.85) y (3.86) por medio de una deformación (véase el principio de deformación de contornos) de los círculos $C_1$ y $C_2$ en cualquier contorno $C$ que se encuentra dentro del anillo $R_2 \leq |z-a| \leq R_1$.

Se puede comprobar que la representación de $f(z)$ mediante la serie de Laurent, ecuación (3.84), es única para un anillo de convergencia dado. Sin embargo, $f(z)$ puede tener diferentes series de Laurent en dos anillos con el mismo centro si entre esos anillos se encuentra cuando menos un punto singular de $f(z)$. La unicidad de la serie de Laurent a menudo nos permite hallar la serie sin evaluar integrales que definen los coeficientes $a_k$.

**Ejemplo 3.23.** *Sea $f(z) = ze^{1/z}$. La función $e^{1/z}$ es analítica en el dominio $|z| > 0$ y, por tanto, usando la serie de Taylor $e^w = 1 + w + w^2/2! + \cdots$ para $w = 1/z$, se obtiene la serie de Laurent de $f(z)$:*

$$f(z) = ze^{1/z} = z\left(1 + \frac{1}{z} + \frac{1}{z^2 2!} + \cdots\right) = z + 1 + \frac{1}{z2!} + \frac{1}{z^2 3!} + \cdots$$

**Ejemplo 3.24.** *Sea $f(z) = 1/(z-1)$. Para ésta, en el dominio $|z| > 1$ se tiene la serie de Laurent:*

$$\begin{aligned} \frac{1}{z-1} &= \frac{1}{z}\frac{1}{1-1/z} = \frac{1}{z}\left(1 + \frac{1}{z} + \frac{1}{z^2} + \cdots\right) \\ &= \frac{1}{z} + \frac{1}{z^2} + \frac{1}{z^3} + \cdots \end{aligned}$$

*En el dominio $|z| < 1$ la función $f(z) = 1/(z-1)$ es analítica y su serie de Laurent se convierte en la serie de Taylor*

$$\frac{1}{z-1} = -\frac{1}{1-z} = -\left(1 + z + z^2 + z^3 + \cdots\right)$$

### 3.3.7. Puntos singulares y cálculo de residuos

Si $z = a$ es un punto singular de una función $f(z)$ y en una vecindad de este punto no hay otros puntos singulares de $f(z)$, entonces el punto $a$ es un **punto singular aislado**.

**Ejemplo 3.25.** *La función $f(z) = 1/z$ es analítica para $|z| > 0$; el punto $z = 0$ es el punto singular aislado.*

**Ejemplo 3.26.** *La función*

$$f(z)=\frac{z-1}{z(z^2+1)}$$

*tiene tres puntos singulares aislados:* $z=0,\ z=i,\ z=-i.$

**Ejemplo 3.27.** *La función*

$$f(z)=\frac{1}{\operatorname{sen}(1/z)}$$

*tiene singularidades en los puntos donde* $\operatorname{sen}(1/z)=0$*; estos puntos son* $a_k=1/\pi k$*,* $k=0,\pm1,\pm2,\ldots$ *Los puntos* $a_k$ *para números finitos* $k=\pm1,\pm2,\ldots$ *son puntos singulares aislados. Sin embargo, el punto singular* $z=0$ *no es aislado, porque en cualquier pequeña vecindad de éste se encuentra un número infinito de puntos singulares aislados,* $\lim_{k\to\infty}a_k=0$.

Si $z=a$ es un punto singular aislado de $f(z)$, entonces, en una vecindad de $a$, la función $f(z)$ puede representarse mediante la serie de Laurent

$$f(z)=\sum_{k=0}^{\infty}a_k(z-a)^k+\sum_{k=1}^{\infty}\frac{a_{-k}}{(z-a)^k}$$

Pueden presentarse dos casos que se consideran a continuación.

**Caso 1.** El número de potencias negativas de $(z-a)$ es finito en la expansión de Laurent,

$$f(z)=\sum_{k=0}^{\infty}a_k(z-a)^k+\sum_{k=1}^{m}\frac{a_{-k}}{(z-a)^k} \tag{3.90}$$

Entonces, se dice que la singularidad de $f(z)$ en el punto $z=a$ es el **polo de orden** $m$. Cuando $m=1$, el polo se llama simple. Una función que es analítica en todo punto en el plano finito recibe el nombre de **función entera**. Una función analítica cuyas únicas singularidades en el plano finito son polos se conoce como **función meromorfa**. Notamos que cuando $f(z)$ tiene un polo de orden $m$, podemos definir la función

$$\begin{aligned}\phi(z)&=(z-a)^m f(z)\quad\text{si}\quad z\neq a\\ \phi(a)&=\lim_{z\to a}(z-a)^m f(z)=a_{-m}\neq0\end{aligned}$$

que es analítica en $z=a$, mientras tanto la función $(z-a)^{m-1}f(z)$ no la es en $z=a$. Esta propiedad puede usarse como la definición de polo de orden $m$.

**Caso 2.** La expansión de Laurent contiene un número infinito de potencias negativas de $(z-a)$, entonces $z=a$ se llama **punto singular esencial** de $f(z)$. El comportamiento de $f(z)$ en la vecindad de un polo y de un punto singular esencial difiere radicalmente, como se verá al final de la sección.

El coeficiente $a_{-1}$ en la serie de Laurent de $f(z)$ en la vecindad de un punto singular aislado $a$ juega un papel importante en la evaluación de las integrales de funciones analíticas. Este coeficiente recibe el nombre de **residuo** de $f(z)$ en el punto $z = a$ y se utiliza la notación

$$a_{-1}(a) = \operatorname{Res} f(z)|_{z=a}$$

Cuando la singularidad en $z = a$ es un polo de orden $m$, el residuo en este punto se puede calcular sin obtener la serie de Laurent. Multiplicando la ecuación (3.90) por $(z-a)^m$ se tiene

$$\begin{aligned} \phi(z) &= (z-a)^m f(z) && (3.91)\\ &= a_{-m} + a_{-m+1}(z-a) + \cdots + a_{-1}(z-a)^{m-1} + a_0(z-a)^m + \cdots \end{aligned}$$

donde $a_{-m} \neq 0$. De la última ecuación se ve que el residuo de una función $f(z)$ en un polo de orden $m$ se calcula mediante la fórmula

$$a_{-1} = \frac{1}{(m-1)!}\frac{d^{m-1}(z-a)^m f(z)}{dz^{m-1}}|_{z=a} \tag{3.92}$$

Para un polo simple, de la ecuación (3.91) o (3.92) se tiene

$$a_{-1} = \lim_{z\to a}(z-a)f(z) \tag{3.93}$$

De ésta se obtiene la fórmula más útil para calcular el residuo en un polo simple. Si $f(z)$ tiene un polo simple en $z = a$, ésta puede presentarse en forma

$$f(z) = \frac{p(z)}{q(z)}$$

donde $p(z)$ y $q(z)$ son analíticas en $z = a$, $p(a) \neq 0$ y $q(z)$ tiene un cero simple en $z = a$. Como consecuencia, $q(z)$ puede desarrollarse en una serie de Taylor de la forma

$$q(z) = (z-a)q'(a) + \frac{(z-a)^2}{2!}q''(a) + \cdots$$

De donde, por la ecuación (3.93),

$$a_{-1} = \lim_{z\to a}(z-a)f(z) = \lim_{z\to a}(z-a)\frac{p(z)}{q(z)} = \frac{p(a)}{q'(a)} \tag{3.94}$$

**Ejemplo 3.28.** *La función*

$$f(z) = \frac{1+z}{z(2-z)}$$

*tiene singularidades en los puntos $z = 0$ y $z = 2$. Notando que la función $\phi(z) = zf(z)$ ya no tiene singularidad en $z = 0$, concluimos que $z = 0$ es polo simple de $f(z)$ y, por tanto, el residuo en este punto es*

$$a_{-1}(0) = \lim_{z\to 0} zf(z) = 1/2$$

*La función $\phi(z) = (z-2) f(z)$ no es singular en $z = 2$; por tanto, este punto es, también, polo simple. El residuo correspondiente es*

$$a_{-1}(2) = \lim_{z\to 2} (z-2) f(z) = -3/2$$

**Ejemplo 3.29.** *La función*

$$f(z) = \frac{e^z}{z^2+1} = \frac{e^z}{(z+i)(z-i)}$$

*tiene polos simples en $z = -i$ y $z = i$. Por tanto, el residuo en $z = -i$ es*

$$a_{-1}(-i) = \lim_{z\to -i} (z+i) f(z) = -e^{-i}/2i$$

*De manera similar se obtiene que*

$$a_{-1}(+i) = \lim_{z\to i} (z-i) f(z) = e^{i}/2i$$

**Ejemplo 3.30.** *La función*

$$f(z) = \frac{1}{z(z+1)^2}$$

*tiene un polo simple en $z = 0$ dado que $\phi(z) = zf(z)$ es analítica en este punto y $\phi(0) \neq 0$. El residuo correspondiente es*

$$a_{-1}(0) = \lim_{z\to 0} zf(z) = 1$$

*Notando que $\phi(z) = (z+1)^2 f(z)$ es analítica en $z = -1$ y $\phi(-1) \neq 0$, concluimos que el punto $z = -1$ es polo de orden dos y el residuo correspondiente es entonces*

$$a_{-1}(-1) = \lim_{z\to -1} \frac{d(z^2+1) f(z)}{dz} = \lim_{z\to -1} \frac{d(1/z)}{dz} = -1$$

**Ejemplo 3.31.** *La función*

$$f(z) = \frac{\operatorname{sen} z}{z^4}$$

*tiene un polo de orden tres en $z = 0$, porque la serie de Laurent de esta función para $|z| > 0$ es*

$$\begin{aligned} f(z) &= \frac{1}{z^4}\left(z - \frac{z^3}{3!} + \frac{z^5}{5!} - \cdots\right) \\ &= \frac{1}{z^3} - \frac{1}{z3!} + \frac{z}{5!} - \ldots \end{aligned}$$

*Calculando el residuo según la fórmula (3.92), se obtiene $a_{-1}(0) = -1/3$.*

**Ejemplo 3.32.** *La función*

$$f(z) = \cos\frac{1}{z-1}$$

*tiene un punto singular aislado en* $z = 1$. *Usando la serie de Taylor de la función*

$$\cos w = 1 - \frac{w^2}{2!} + \frac{w^4}{4!} - \cdots$$

*y haciendo* $w = 1/(z-1)$, *se obtiene la expansión de Laurent,*

$$\cos\frac{1}{z-1} = 1 - \frac{1}{2!\,(z-1)^2} + \frac{1}{4!\,(z-1)^4} - \cdots$$

*Dado que la serie de Laurent tiene un número infinito de potencias negativas de* $(z-1)$, *el punto singular aislado* $z = 1$ *es esencial. Sin embargo, el residuo en este punto es cero,* $a_{-1}(1) = 0$, *puesto que el término* $(z-1)^{-1}$ *no aparece en la serie de Laurent.*

**Ejemplo 3.33.** *La función*

$$f(z) = \exp(1/z)$$

*tiene una singularidad esencial en* $z = 0$. *La serie de Laurent,*

$$\exp(1/z) = 1 + \frac{1}{z} + \frac{1}{2!z^2} + \frac{1}{3!z^3} + \cdots$$

*tiene todas las potencias negativas de* $z$. *Nótese que la* $f(z) = \exp(1/z)$ *no tiene límite si* $z$ *tiende a cero a lo largo del eje imaginario; se hace infinita si* $z \to 0$ *a lo largo del eje real positivo, pero tiende a cero si* $z \to 0$ *a lo largo del eje real negativo. Mostramos que dicha función toma cualquier valor dado* $c = \lambda e^{i\alpha} = \lambda e^{i(\alpha+2\pi k)} \neq 0$ *en una vecindad arbitrariamente pequeña de* $z = 0$, *donde* $k$ *es un número natural. En efecto, suponiendo* $z = \rho(\cos\varphi + i\,\mathrm{sen}\,\varphi) = \rho\exp(i\varphi)$, *debe resolverse la ecuación*

$$\exp(1/z) = \exp((\cos\varphi - i\,\mathrm{sen}\,\varphi)/\rho) = \lambda e^{i(\alpha+2\pi k)}$$

*con respecto a* $\rho$ *y* $\varphi$. *Igualando los valores absolutos y argumentos, se tiene*

$$\cos\varphi = \rho\ln\lambda, \quad \mathrm{sen}\,\varphi = -\rho(\alpha + 2\pi k)$$

*A partir de estas ecuaciones se obtienen las soluciones*

$$\rho^2 = \frac{1}{(\ln\lambda)^2 + (\alpha+2\pi k)^2}, \quad \tan\varphi = -\frac{\alpha+2\pi k}{\ln\lambda}$$

*De donde* $\rho$ *puede hacerse arbitrariamente pequeño al tomar el número* $k$ *suficientemente grande, dejando el número* $c$ *inalterado.*

El ejemplo anterior ilustra el **teorema de Picard**, la demostración de éste es omitida por ser un tanto complicada.

**Teorema 3.7.** *Si una función analítica $f(z)$ tiene una singularidad esencial aislada en un punto $a$, ésta toma cualquier valor dado un número infinito de veces en una vecindad arbitrariamente pequeña de $a$.*

Hasta el momento nos restringimos a la consideración de funciones analíticas univaluadas. Cuando surgen funciones multivaluadas, podemos considerar sus ramas univaluadas. Estas ramas pueden tener derivadas en ciertos dominios del plano complejo $z$ y, por lo tanto, son analíticas en estos dominios. Por ejemplo, la función $f(z) = \sqrt{z} = \rho^{1/2}\left[\cos\left((\theta + 2\pi k)/2\right) + i\,\mathrm{sen}\left((\theta + 2\pi k)/2\right)\right]$, $k = 0, 1$, tiene dos ramas univaluadas y cada una de ellas tiene la derivada $f'(z) = 1/\left(2\sqrt{z}\right)$. Por lo tanto, cada una de estas ramas es analítica, excepto en el punto $z = 0$. Otro ejemplo es la función $f(z) = \ln z = u + iv = \ln \rho + i(\theta + 2\pi k)$, $k = 0, \pm 1, \pm 2, \ldots$, que tiene un número infinito de ramas, cada una de ellas tiene la derivada $f'(z) = 1/z$. El punto $z = 0$ es el único en donde la derivada no existe y todas las ramas dejan de ser analíticas en éste. Los puntos singulares de los dos ejemplos no son del tipo puntos aislados, dado que cada rama de $\sqrt{z}$ y $\ln z$ es discontinua en la vecindad de su punto singular. Sin profundizar el tema, notamos que esos puntos singulares se llaman **puntos de ramificación** y no son puntos singulares aislados de $f(z)$. En la vecindad de su punto de ramificación, $f(z)$ no puede representarse mediante la serie de Laurent.

### 3.3.8. Evaluación de integrales por el método de residuos

Los residuos de las funciones analíticas son sumamente útiles para la evaluación de integrales de contorno y ciertas integrales reales de manera elegante y sencilla.

**Teorema 3.8.** *Sea $f(z)$ una función analítica en el interior de una trayectoria simple cerrada $C$ y sobre $C$, excepto un número finito de puntos singulares aislados $a_1, a_2, \ldots, a_n$ en el interior de $C$. Entonces, la integral de $f(z)$ a lo largo de $C$ es igual a $2\pi i$ multiplicado por la suma de residuos en dichos puntos singulares,*

$$\oint_C f(z)\,dz = 2\pi i \sum_{k=1}^{n} Res\, f(z)\big|_{z=a_k} \tag{3.95}$$

*donde la integral se toma en sentido contrario al de las manecillas del reloj.*

**Demostración.** *Se encierra cada uno de los puntos singulares por un círculo suficientemente pequeño $C_k$, de tal manera que $f(z)$ sea analítica en el dominio múltiplemente conexo limitado por $C$ y los $C_k$. Por el teorema integral de Cauchy, ecuación (3.67), se tiene*

$$\oint_C f(z)\,dz = \sum_{k=1}^{n} \oint_{C_k} f(z)\,dz$$

*Puesto que, por la ecuación (3.86) para $k = 1$,*

$$\frac{1}{2\pi i}\oint_{C_k} f(z)\,dz = a_{-1}(a_k) = Res\, f(z)|_{z=a_k}$$

*la fórmula (3.67) conduce a la (3.95) y el teorema queda demostrado.* ■

A continuación presentamos algunos ejemplos de aplicación del teorema de residuos en relación con las integrales complejas y reales.

**Ejemplo 3.34.** *La función*

$$f(z) = \frac{1+z}{z(2-z)}$$

*tiene dos polos simples en $z = 0$ y $z = 2$ con residuos correspondientes $a_{-1}(0) = 1/2$ y $a_{-1}(2) = -3/2$. Por lo tanto, la integral de esta función por una trayectoria simple cerrada $C$ que: 1) encierre el punto $z = 0$ es igual a $\pi i$; 2) encierre el punto $z = 2$ es igual a $-3\pi i$; y 3) encierre los dos puntos a la vez es igual a $2\pi i\,(1/2 - 3/2) = -2\pi i$.*

**Ejemplo 3.35.** *Para evaluar la integral*

$$\oint_C f(z)\,dz = \oint_C \frac{e^z}{z^2+1}dz$$

*por el círculo $|z| = 2$, notamos que el integrando tiene dos polos simples en el interior de la trayectoria, $z = i$ y $z = -i$. Los residuos correspondientes son $a_{-1}(i) = e^i/2i$ y $a_{-1}(-i) = -e^{-i}/2i$. Por tanto, la integral es igual a $2\pi i\,(e^i/2i - e^{-i}/2i) = 2\pi i \operatorname{sen} 1$.*

**Ejemplo 3.36.** *En la sección anterior encontramos que $f(z) = \cos(1/(z-1))$ tiene una singularidad esencial en el punto $z = 1$. Sin embargo, el residuo correspondiente es cero, $a_{-1}(1) = 0$. Por tanto, el valor de integral de esta función es cero para todo contorno $C$ que no pasa por el punto $z = 1$. Si $z = 1$ se encuentra en $C$, entonces la integral es impropia y requiere otros métodos para evaluarla.*

**Ejemplo 3.37.** *La función $f(z) = 1/\left(z^3-1\right)^2$ tiene tres polos de orden dos en los puntos $z = 1$, $z = e^{i2\pi/3}$ y $z = e^{i4\pi/3}$. El contorno $|z-1| = 1$ contiene sólo el polo $z = 1$. Entonces, utilizando la ecuación (3.92), la integral de dicha función a lo largo del círculo $|z-1| = 1$ es igual a $2\pi i a_{-1}(1) = 2\pi i\,(-2/9) = -4\pi i/9$.*

El teorema de residuos proporciona también un método magnífico para evaluar ciertas clases de integrales reales complicadas. En primer lugar se consideran integrales del tipo

$$I = \int_0^{2\pi} R(\cos\theta, \operatorname{sen}\theta)\,d\theta \tag{3.96}$$

donde $R(\cos\theta, \operatorname{sen}\theta)$ es una función racional real de $\cos\theta$ y $\operatorname{sen}\theta$ finita sobre el intervalo $0 \leq \theta \leq 2\pi$. Haciendo $z = e^{i\theta}$, se tiene

$$\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2} = \frac{1}{2}\left(z + \frac{1}{z}\right), \quad \operatorname{sen}\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i} = \frac{1}{2i}\left(z - \frac{1}{z}\right)$$

De esta manera, el integrando se transforma en una función racional de $z$, es decir, $f(z)$. Como $\theta$ varía desde 0 hasta $2\pi$, la variable $z$ recorre una vez la trayectoria alrededor del círculo unitario $|z| = 1$ en sentido contrario al de las manecillas del reloj. Se tiene $dz/d\theta = ie^{i\theta} = iz$. Por tanto, $d\theta = dz/iz$ y la integral dada toma la forma

$$I = \oint_C \frac{f(z)}{iz} dz = 2\pi i \sum \operatorname{Res}\left(\frac{f(z)}{iz}\right) \tag{3.97}$$

donde los residuos se calculan en los puntos singulares de $(f(z)/iz)$ en el interior del círculo unitario $C$.

**Ejemplo 3.38.** *Sea $p$ un número real fijo en el intervalo $0 < p < 1$. Se considera*

$$\begin{aligned} I &= \int_0^{2\pi} \frac{d\theta}{1 - 2p\cos\theta + p^2} = \oint_C \frac{dz/iz}{1 - p(z + 1/z) + p^2} \\ &= \oint_C \frac{dz}{i(1 - pz)(z - p)} \end{aligned}$$

*El integrando tiene dos polos simples en $z = 1/p$ y $z = p$. Sólo el último polo se encuentra en el interior del círculo unitario $C$, y el residuo es*

$$Res \left. \frac{1}{i(1 - pz)(z - p)} \right|_{z=p} = \frac{1}{i(1 - p^2)}$$

*Entonces, la integral dada es $I = 2\pi/(1 - p^2)$.*

Otro tipo de integrales que se prestan para evaluarlas por el teorema de residuos son las integrales impropias. Una integral así es la que tiene el intervalo de integración infinito

$$\int_{-\infty}^{\infty} f(x)\,dx = \lim_{A\to\infty} \int_{-A}^{0} f(x)\,dx + \lim_{B\to\infty} \int_0^B f(x)\,dx$$

siendo $f(x)$ una función real finita en el intervalo de integración. Si los dos límites existen, la integral puede escribirse en la forma

$$\int_{-\infty}^{\infty} f(x)\,dx = \lim_{A\to\infty} \int_{-A}^{A} f(x)\,dx.$$

Suponemos que $f(x)$ es tal que su homólogo complejo $f(z)$ tiene el crecimiento no más rápido que $O\left(r^{-1-\delta}\right)$ cuando $|z| = r$ es suficientemente grande, es

decir, $|f(z)| < Mr^{-1-\delta}$, donde $M$ es una constante y $\delta > 0$. En este caso se considera la integral de contorno $C$ que consiste del intervalo $(-A, A)$ y la semicircunferencia $S$ que une los puntos $x = A$ y $x = -A$ en el semiplano superior,

$$\oint_C f(z)\,dz = \int_{-A}^{A} f(x)\,dx + \int_S f(z)\,dz = 2\pi i \sum \operatorname{Res} f(z)$$

donde la suma se extiende a todos los residuos de $f(z)$ en los puntos singulares aislados que se encuentran en el interior del contorno $C$. A partir de esto, se tiene que

$$\int_{-\infty}^{\infty} f(x)\,dx = \lim_{A\to\infty} \int_{-A}^{A} f(x)\,dx = 2\pi i \sum \operatorname{Res} f(z) \tag{3.98}$$

dado que

$$\left|\int_S f(z)\,dz\right| < Mr^{-1-\delta}\pi r = M\pi r^{-\delta} \to_{r\to\infty} 0$$

En la fórmula (3.98) la suma se extiende a todos los residuos de $f(z)$ en los puntos singulares aislados que se encuentran en el semiplano superior.

**Ejemplo 3.39.** *Para evaluar la integral*

$$I = \int_0^{\infty} \frac{dx}{1+x^4}$$

*notamos que la función* $f(x) = \left(1+x^4\right)^{-1}$ *es par y, además, se cumple la condición* $\left|\left(1+z^4\right)^{-1}\right| < Mr^{-4}$. *Por tanto,*

$$I = \frac{1}{2}\int_{-\infty}^{\infty} \frac{dx}{1+x^4} = 2\pi i \sum Res\, f(z)$$

*La función* $f(z) = \left(1+z^4\right)^{-1}$ *tiene cuatro polos simples en los puntos* $a_1 = e^{i\pi/4}$, $a_2 = e^{i3\pi/4}$, $a_3 = e^{i5\pi/4}$, $a_4 = e^{i7\pi/4}$, *pero sólo los dos primeros de estos polos se encuentran en el semiplano superior. Por la ecuación (3.94), los residuos correspondientes son*

$$\begin{aligned} a_{-1}(a_1) &= \frac{1}{(1+z^4)'}\,|_{z=a_1} = \frac{1}{4z^3}\,|_{z=a_1} = -\frac{1}{4}e^{i\pi/4} \\ a_{-1}(a_2) &= \frac{1}{(1+z^4)'}\,|_{z=a_2} = \frac{1}{4z^3}\,|_{z=a_2} = \frac{1}{4}e^{-i\pi/4} \end{aligned}$$

*Finalmente,*

$$I = \frac{1}{2}\int_{-\infty}^{\infty} \frac{dx}{1+x^4} = 2\pi i \frac{1}{2} \sum Res\, f(z) = \frac{\pi}{2}\operatorname{sen}(\pi/4) = \frac{\pi}{2\sqrt{2}}$$

La fórmula (3.98) se extiende fácilmente a las integrales reales que se presentan en relación con la integral de Fourier,

$$F_c(\omega) = \int_{-\infty}^{\infty} f(x) \cos \omega x dx, \quad F_s(\omega) = \int_{-\infty}^{\infty} f(x) \operatorname{sen} \omega x dx \tag{3.99}$$

donde $\omega$ es un número real. Si $f(x)$ satisface a la hipótesis que $|f(z)| \sim O\left(r^{-1-\delta}\right)$ con $\delta > 0$ cuando $|z| = r$ es suficientemente grande, entonces, las integrales $F_c(\omega)$ y $F_s(\omega)$ pueden evaluarse de manera semejante a las de la ecuación (3.98). Considérese la integral auxiliar

$$\oint_C f(z) e^{i\omega z} dz, \quad \omega > 0 \tag{3.100}$$

sobre el contorno $C$ que consiste del intervalo $(-A, A)$ y la semicircunferencia $S$ que une los puntos $x = A$ y $x = -A$ en el semiplano superior. Como $S$ se encuentra en el semiplano superior, $y \geq 0$ y $\omega > 0$, se ve que

$$\left|f(z) e^{i\omega z}\right| = |f(z)| \left|e^{i\omega x}\right| \left|e^{-\omega y}\right| \leq |f(z)| \sim O\left(r^{-1-\delta}\right)$$

y el valor de la integral sobre $S$ tiende a cero conforme $r$ tiende al infinito. Luego, en lugar de la ecuación (3.98), se obtiene la fórmula

$$F(\omega) = \int_{-\infty}^{\infty} f(x) e^{i\omega x} dx = \lim_{A \to \infty} \int_{-A}^{A} f(x) e^{i\omega x} dx = 2\pi i \sum \operatorname{Res}\left[f(z) e^{i\omega z}\right] \tag{3.101}$$

donde se suman los residuos de $f(z) e^{i\omega z}$ en sus puntos singulares del semiplano superior. De la ecuación (3.101) se obtiene

$$\begin{aligned} F_c(\omega) &= \int_{-\infty}^{\infty} f(x) \cos \omega x dx = \operatorname{Re} F(\omega) \\ &= -2\pi \sum \operatorname{Im}\left(\operatorname{Res}\left[f(z) e^{i\omega z}\right]\right) \qquad (3.102) \\ F_s(\omega) &= \int_{-\infty}^{\infty} f(x) \operatorname{sen} \omega x dx = \operatorname{Im} F(\omega) \\ &= 2\pi \sum \operatorname{Re}\left(\operatorname{Res}\left[f(z) e^{i\omega z}\right]\right) \qquad (3.103) \end{aligned}$$

para $\omega > 0$.

**Ejemplo 3.40.** *Demuéstrese que*

$$I_1 = \int_{-\infty}^{\infty} \frac{\cos \omega x}{k^2 + x^2} dx = \frac{\pi}{k} e^{-k\omega}, \quad I_2 = \int_{-\infty}^{\infty} \frac{\operatorname{sen} \omega x}{k^2 + x^2} dx = 0$$

*En efecto, la función $e^{i\omega z} / \left(k^2 + z^2\right)$ tiene sólo un polo en el semiplano superior, a saber, un polo simple en $z = ik$. El residuo correspondiente es*

$$Res \left.\frac{e^{i\omega z}}{k^2 + z^2}\right|_{z=ik} = \frac{e^{i\omega z}}{2z} \,|_{z=ik} = \frac{e^{-\omega k}}{2ik}$$

*Por lo tanto,*

$$I_1 = -2\pi \operatorname{Im}\left(\frac{e^{-\omega k}}{2ik}\right) = \frac{\pi}{k}e^{-k\omega}, \quad I_2 = 2\pi \operatorname{Re}\left(\frac{e^{-\omega k}}{2ik}\right) = 0$$

*Nótese que el integrando de $I_2$ es una función impar y el intervalo de integración es simétrico con respecto a cero. Por tanto, la integral $I_2$ debe ser igual a cero, que confirma nuestro resultado anterior.*

Cabe mencionar que existen más clases de integrales reales que pueden evaluarse aplicando el teorema de residuos a integrales complejas apropiadas.

**Nota bibliográfica:** para ampliar los conocimientos sobre el tema de este capítulo, se puede consultar [1], [4], [6], [12], [15] y [17].

# Capítulo 4

# Geometría analítica y diferencial

## 4.1. Curvas en un plano: líneas rectas y secciones cónicas

### 4.1.1. Representación de curvas en 2-D

El estudio de las curvas y superficies en el espacio, por medio del cálculo, es una rama importante de las Matemáticas, conocida como **geometría diferencial**. Antes de pasar a los conceptos fundamentales de la geometría diferencial, nos refrescaremos la memoria con algunos hechos básicos acerca de las curvas en el plano. Las curvas se presentan en muchos problemas de Cálculo, así como de Física, principalmente como trayectorias de movimiento.

Una **curva** en el espacio es un objeto geométrico unidimensional, es decir, se caracteriza por sólo una variable independiente. Una **curva plana** es aquella que se encuentra en un plano en el espacio; si una curva no es plana se dice que es una **curva alabeada**. Una vez dado un sistema de coordenadas cartesianas, se puede representar una curva $C$ en el plano mediante una función vectorial

$$\mathbf{r}\,(t) = x\,(t)\,\mathbf{i} + y\,(t)\,\mathbf{j} \tag{4.1}$$

donde a cada valor de la variable real $t$ le corresponde un punto de $C$ que tiene el vector de posición con coordenadas $(x\,(t)\,, y\,(t))$. Una representación así se llama representación paramétrica de $C$ y $t$ es el parámetro de ésta (véase la figura 4.1). En forma escalar, la ecuación (4.1) se representa mediante las dos ecuaciones

$$x = x\,(t)\,, \quad y = y\,(t) \tag{4.2}$$

Otros tipos de representaciones de curvas planas son

$$y = f\,(x) \tag{4.3}$$

$$\varphi\,(x, y) = 0 \tag{4.4}$$

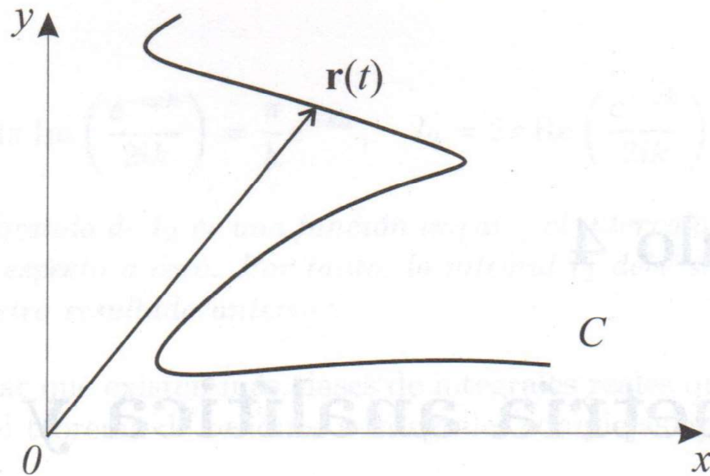



Figura 4.1. Representación paramétrica de una curva plana

que establecen relaciones funcionales entre coordenadas de puntos en una curva. Las curvas correspondientes a las ecuaciones $\varphi(x,y)=0$ y $\lambda\varphi(x,y)=0$, donde $\lambda \neq 0$ es una constante, son idénticas.

Los pares de coordenadas $(x,y)$ que satisfacen simultáneamente a las ecuaciones de dos curvas,

$$\varphi_1(x,y)=0, \quad \varphi_2(x,y)=0 \tag{4.5}$$

representan los puntos de intersección de estas curvas. En particular, si la ecuación $\varphi(x,0)=0$, que es equivalente a las dos ecuaciones $\varphi(x,y)=0$ y $y=0$, tiene una o más raíces reales, éstas son abscisas de intersección de la curva $\varphi(x,y)=0$ con el eje $x$.

Para cualquier número real $\lambda$, la ecuación

$$\varphi_1(x,y)+\lambda\varphi_2(x,y)=0 \tag{4.6}$$

describe una curva que pasa por todos los puntos de intersección de las dos curvas, $\varphi_1=0$ y $\varphi_2=0$.

La unión de las curvas, $\varphi_1=0$ y $\varphi_2=0$, es representada por la ecuación

$$\varphi_1(x,y)\cdot\varphi_2(x,y)=0 \tag{4.7}$$

### 4.1.2. Líneas rectas

Cualquier **recta** $L$ en el plano se puede representar en forma paramétrica

$$\mathbf{r}(t)=\mathbf{r}_1+\mathbf{a}t \tag{4.8}$$

donde $\mathbf{a}$ y $\mathbf{r}_1$ son vectores constantes. Esta recta pasa por el punto con el vector de posición $\mathbf{r}_1$ y tiene la dirección de $\mathbf{a}$ (véase la figura 4.2). Si el vector direccional es unitario, $|\mathbf{a}|=1$, sus componentes son los cosenos directores de $L$. En este caso $|t|$ mide la distancia de los puntos de $L$ al punto $\mathbf{r}_1$. En forma escalar paramétrica $L$ se describe por medio de las dos ecuaciones $x(t)=x_1+a_xt, \quad y(t)=y_1+a_yt$.

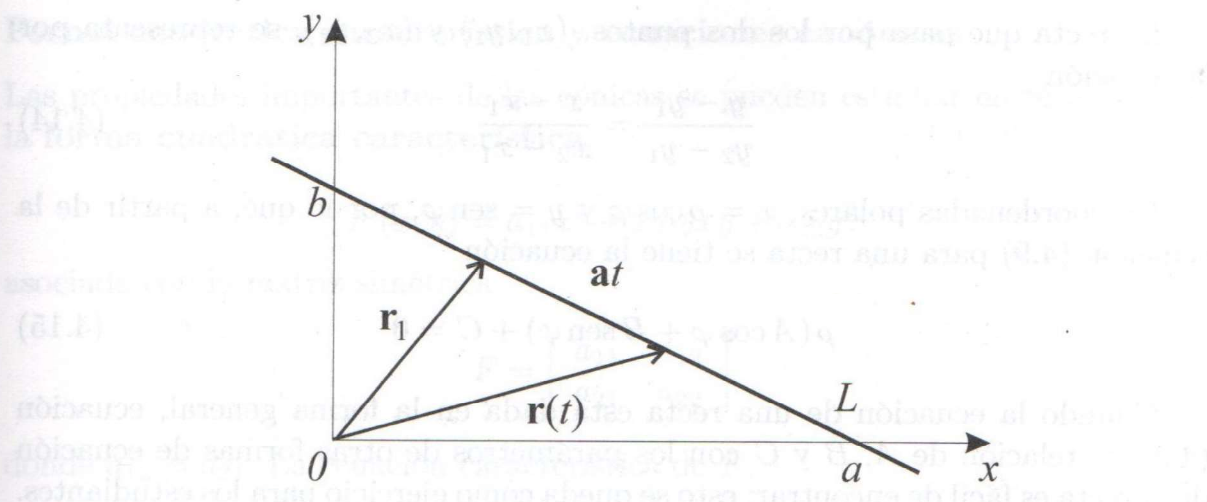



Figura 4.2. Representación paramétrica de una recta

Una vez dado un sistema de coordenadas cartesianas, se puede representar cualquier recta en el plano mediante una ecuación lineal de la forma

$$Ax + By + C = 0 \tag{4.9}$$

donde $A$, $B$ y $C$ son constantes reales. El caso especial $C = 0$ corresponde a una recta que pasa por el centro del sistema de coordenadas.

Las formas especiales de las ecuaciones de una línea recta son como siguen. La ecuación

$$y = mx + b \tag{4.10}$$

representa una recta que intersecta el eje $y$ a la distancia $b$ del centro de coordenadas y tiene la pendiente $m = \tan\theta$, siendo $\theta$ el ángulo entre la recta y el eje $x$.

La recta, definida por la ecuación

$$\frac{x}{a} + \frac{y}{b} - 1 = 0 \tag{4.11}$$

intersecta el eje $x$ a la distancia $a$ y el eje $y$ a la distancia $b$ del centro del sistema de coordenadas.

Forma normal. Sea $p$ la longitud de la perpendicular $\mathbf{p}$ dirigida desde el origen del sistema de coordenadas a la recta y sea $\theta$ el ángulo entre el eje $x$ positivo y la perpendicular medido en sentido opuesto al de las manecillas del reloj. Entonces, la ecuación de la recta es

$$\mathbf{r} \cdot \mathbf{n} = p \quad \text{o} \quad x\cos\theta + y\,\mathrm{sen}\,\theta - p = 0 \tag{4.12}$$

donde $p = |\mathbf{p}|$ y $\mathbf{p} = p\mathbf{n}$.

La línea recta que pasa por el punto $(x_1, y_1)$ y tiene la pendiente $m$ se describe por la ecuación

$$y - y_1 = m\,(x - x_1) \tag{4.13}$$

La recta que pasa por los dos puntos, $(x_1, y_1)$ y $(x_2, y_2)$, se representa por la ecuación

$$\frac{y - y_1}{y_2 - y_1} = \frac{x - x_1}{x_2 - x_1} \tag{4.14}$$

En coordenadas polares, $x = \rho \cos\varphi$ y $y = \operatorname{sen}\varphi$, por lo que, a partir de la ecuación (4.9) para una recta se tiene la ecuación

$$\rho\,(A\cos\varphi + B\operatorname{sen}\varphi) + C = 0 \tag{4.15}$$

Cuando la ecuación de una recta está dada en la forma general, ecuación (4.9), la relación de $A$, $B$ y $C$ con los parámetros de otras formas de ecuación de la recta es fácil de encontrar; esto se queda como ejercicio para los estudiantes.

### 4.1.3. Curvas de segundo orden en 2-D (secciones cónicas)

**Ecuación general de segundo orden**

Las curvas de segundo orden o **secciones cónicas** se representan por la ecuación general de segundo orden

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{13}x + 2a_{23}y + a_{33} = 0 \tag{4.16}$$

o, haciendo $a_{ij} = a_{ji}$ $(i, j = 1, 2, 3)$, se obtiene la ecuación equivalente

$$(a_{11}x + a_{12}y + a_{13})\,x + (a_{21}x + a_{22}y + a_{23})\,y + (a_{31}x + a_{32}y + a_{33}) = 0 \tag{4.17}$$

Resulta que las tres magnitudes

$$I = a_{11} + a_{22}, \quad D = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \tag{4.18}$$

y el signo de

$$\tilde{A} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} \tag{4.19}$$

son invariantes de las ecuaciones (4.16) y (4.17) con respecto a las transformaciones de traslación y rotación. Esos invariantes definen las propiedades de las secciones cónicas y no dependen de la posición de una sección cónica en el plano. Omitiendo la tabla completa, presentamos algunos ejemplos de clasificación de secciones cónicas.

Si $A \neq 0$, $D > 0$ y $A/I < 0$, la sección cónica es una elipse real (véase la figura 4.3).

Si $A \neq 0$ y $D < 0$, la sección cónica es una hipérbola (véase la figura 4.4).

Si $A \neq 0$, $D = 0$ y $\tilde{A} < 0$, la sección cónica correspondiente es una parábola (véase la figura 4.5).

**Forma cuadrática característica y ecuaciones canónicas**

Las propiedades importantes de las cónicas se pueden estudiar en términos de la **forma cuadrática característica**,

$$F(x,y) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2 \tag{4.20}$$

asociada con la matriz simétrica

$$F = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \tag{4.21}$$

donde $a_{12} = a_{21}$. La ecuación característica de $F$,

$$|F - \lambda E| = 0 \quad \text{o} \quad \lambda^2 + \lambda I + D = 0 \tag{4.22}$$

tiene las dos raíces

$$\lambda_{1,2} = \frac{I}{2} \pm \frac{\sqrt{I^2 - 4D}}{2} \tag{4.23}$$

Dado que $I^2 - 4D = (a_{11} - a_{22})^2 + a_{12}^2 > 0$, $\lambda_1$ y $\lambda_2$ son necesariamente reales. Si $D = 0$, entonces, $\lambda_1 = I$ y $\lambda_2 = 0$. Los eigenvectores correspondientes a los eigenvalores $\lambda_1$ y $\lambda_2$, soluciones de la ecuación $F\mathbf{u} = \lambda\mathbf{u}$, además corresponden a los **ejes *(direcciones)* principales** o **ejes de simetría** de la cónica. Así, los ejes principales están direccionados a lo largo de los eigenvectores de $F$. Los cosenos directores, $\cos\theta$ y $\operatorname{sen}\theta$, de la normal a un eje principal satisfacen las ecuaciones

$$\begin{aligned} (a_{11} - \lambda)\cos\theta + a_{12}\operatorname{sen}\theta = 0 \\ a_{21}\cos\theta + (a_{22} - \lambda)\operatorname{sen}\theta = 0 \end{aligned} \tag{4.24}$$

donde $\lambda \neq 0$ es una raíz de la ecuación característica. El ángulo $\vartheta$ entre el eje $x$ positivo y un eje principal de la cónica satisface a la condición

$$\tan 2\vartheta = \tan 2\theta = \frac{2a_{12}}{a_{11} - a_{22}} \tag{4.25}$$

Por medio de una transformación de traslación se pueden eliminar los términos lineales en $x$ y $y$ de la ecuación general de una sección cónica, ecuación (4.16), quedando solamente los términos cuadráticos (la forma cuadrática) y una constante. Esta transformación corresponde al cambio del sistema de coordenadas original por otro que tiene su origen en el centro de la cónica (intersección de los dos ejes principales). Luego, rotando el nuevo sistema de coordenadas el ángulo $\vartheta$, hacemos coincidir los ejes del sistema de coordenadas con los ejes principales de la cónica. En este sistema de coordenadas, la ecuación de una sección cónica toma una de las **formas estándares *(canónicas)*** que están listadas a continuación:

La ecuación canónica de una elipse es

$$\begin{gathered} x^2/a^2 + y^2/b^2 = 1, \quad \text{donde} \\ a^2 = -\tfrac{A}{\lambda_2 D} = -\tfrac{A}{\lambda_1\lambda_2^2}, \quad b^2 = -\tfrac{A}{\lambda_1 D} = -\tfrac{A}{\lambda_1^2\lambda_2} \end{gathered} \tag{4.26}$$

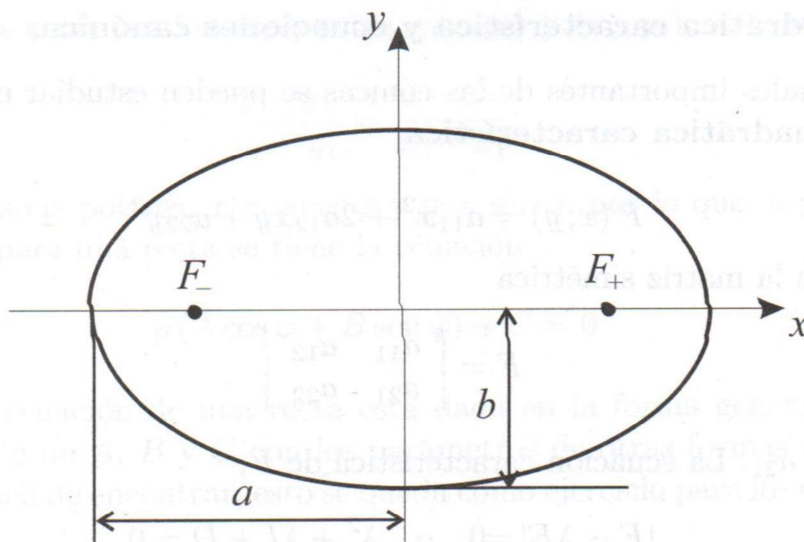



Figura 4.3. Elipse en forma canónica. La excentricidad $\epsilon = \sqrt{1 - b^2/a^2} < 1$, los focos $F_+ = (a\epsilon, 0)$ y $F_- = (-a\epsilon, 0)$

La forma canónica para una hipérbola es

$$x^2/a^2 - y^2/b^2 = 1, \quad \text{donde}$$
$$a^2 = -\tfrac{A}{\lambda_1 D} = -\tfrac{A}{\lambda_1^1 \lambda_2}, \quad b^2 = +\tfrac{A}{\lambda_2 D} = +\tfrac{A}{\lambda_1 \lambda_2^2} \tag{4.27}$$

En forma canónica, una parábola se describe por la ecuación

$$y^2 = 4px, \quad p = \frac{1}{2I}\sqrt{-\frac{A}{I}} == \frac{1}{\lambda_1}\sqrt{-\frac{A}{\lambda_1}}, \quad \lambda_2 = 0 \tag{4.28}$$

Nótese que tanto una elipse como una hipérbola tienen dos ejes principales, mientras que una parábola tiene un solo eje de simetría.

Las ecuaciones canónicas de secciones cónicas degeneradas son como siguen.

Reescribimos la ecuación canónica de una elipse en la forma

$$\frac{x^2}{Ma^2} + \frac{y^2}{Mb^2} = 1$$

donde $M$ puede considerarse como un factor de escalamiento. Multiplicando la ecuación anterior por $M$ y tendiendo $M \to 0$, se obtiene la ecuación de un punto

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 0 \tag{4.29}$$

Este punto se puede considerar como una elipse reducida a la dimensión cero.

De la ecuación de una hipérbola,

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = M$$

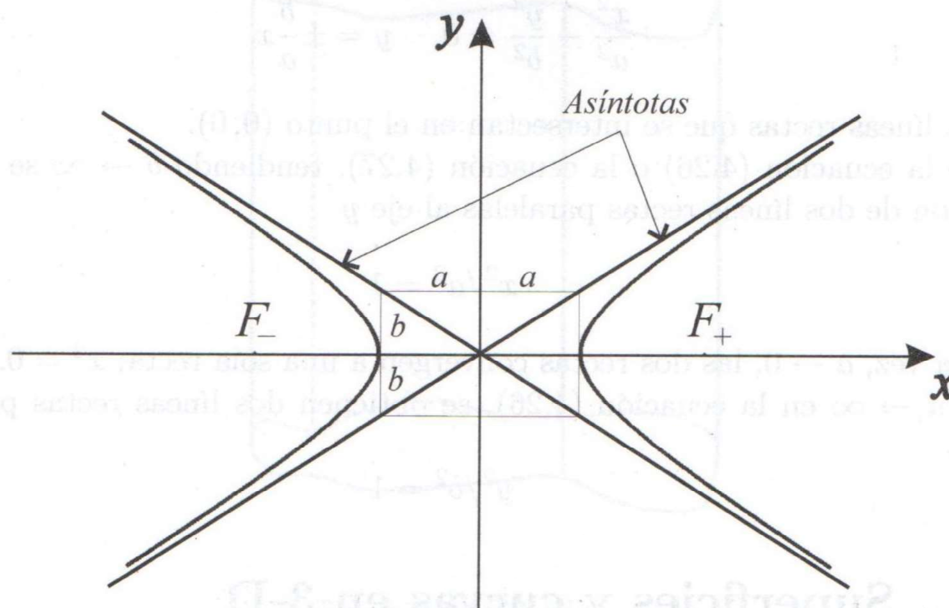



Figura 4.4. Hipérbola en forma canónica. La excentricidad $\epsilon = \sqrt{1 + b^2/a^2} > 1$, los focos $F_+ = (a\epsilon, 0)$ y $F_- = (-a\epsilon, 0)$

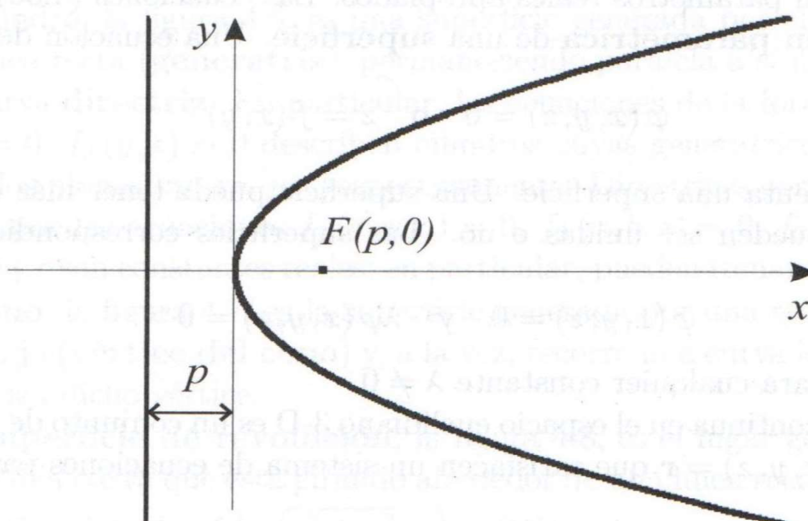



Figura 4.5. Parábola en forma canónica. La excentricidad $\epsilon = 1$, el foco $F = (p, 0)$

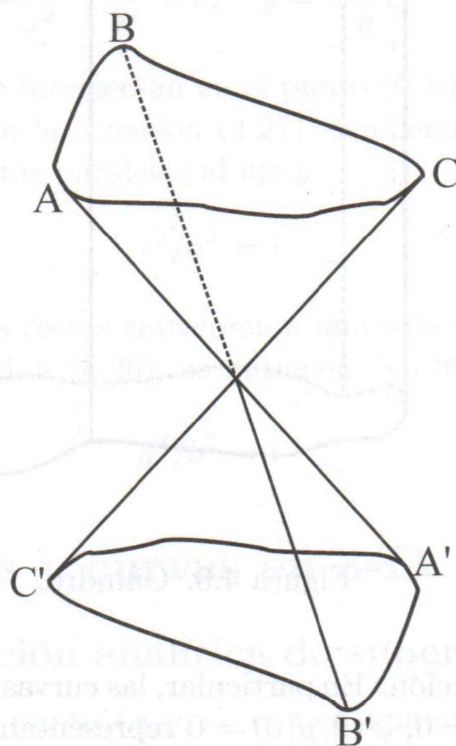



Figura 4.7. Cono

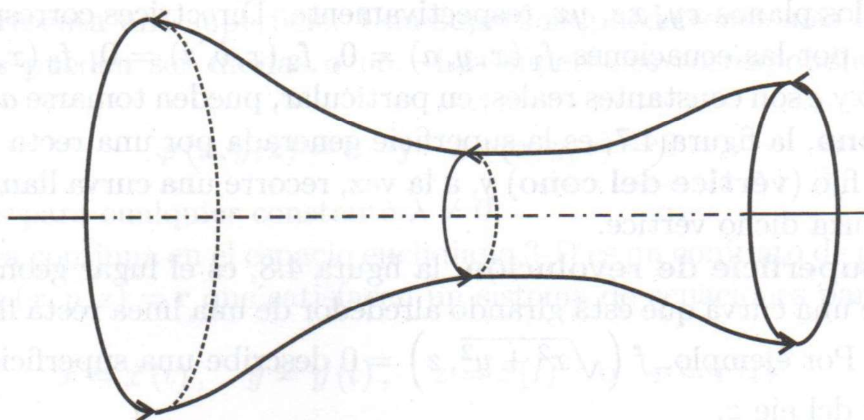


Figura 4.8. Superficie de revolución

La ecuación

$$\varphi_1(x, y, z) \cdot \varphi_2(x, y, z) = 0 \tag{4.38}$$

corresponde a la superficie que es la unión de las dos superficies $\varphi_1 = 0$ y $\varphi_2 = 0$.

La familia de curvas de un parámetro $\lambda$, definida por dos ecuaciones

$$\varphi_1(x, y, z|\lambda) = 0 \quad \text{y} \quad \varphi_2(x, y, z|\lambda) = 0 \tag{4.39}$$

genera, a la vez, una superficie cuya ecuación puede obtenerse eliminando $\lambda$ de las ecuaciones (4.39).

### 4.2.2. Planos

Una vez dado un sistema de coordenadas cartesianas, una ecuación lineal de la forma general en las variables $x, y, z$,

$$Ax + By + Cz + D = 0 \quad \text{o} \quad \mathbf{A} \cdot \mathbf{r} + D = 0 \tag{4.40}$$

donde $A$, $B$ y $C$ son constantes que no se hacen igual a cero simultáneamente, representa un **plano** en el espacio de 3-D, y viceversa, todo plano en el espacio de 3-D puede representarse por la ecuación (4.40). Las constantes $A$, $B$ y $C$ son números directores de la normal a este plano; el vector $\mathbf{A} = (A, B, C)$ tiene la dirección de la normal. El caso especial $D = 0$ corresponde a un plano que pasa por el origen del sistema de coordenadas.

Otras formas de la ecuación de un plano son como siguen.

El plano que intersecta al eje $x$ en $x = a$, al eje $y$ en $y = b$ y al eje $z$ en $z = c$ se representa por la ecuación

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} - 1 = 0 \tag{4.41}$$

Sea $\mathbf{n} = (\cos\alpha_x,\ \cos\alpha_y,\ \cos\alpha_z)$ un vector normal unitario a un plano y sea $\mathbf{r}_1 = (x_1, y_1, z_1)$ el vector de un punto fijo de este plano. Entonces, la ecuación de dicho plano es

$$(\mathbf{r} - \mathbf{r}_1) \cdot \mathbf{n} = 0 \quad \text{o} \tag{4.42}$$

$$(x - x_1)\cos\alpha_x + (y - y_1)\cos\alpha_y + (z - z_1)\cos\alpha_z = 0 \tag{4.43}$$

De la última ecuación se tiene

$$x\cos\alpha_x + y\cos\alpha_y + z\cos\alpha_z - p = 0 \tag{4.44}$$

donde $p = x_1\cos\alpha_x + y_1\cos\alpha_y + z_1\cos\alpha_z$ es la longitud de la normal desde el origen hasta el plano.

Si se conocen tres puntos $\mathbf{r}_1 = (x_1, y_1, z_1)$, $\mathbf{r}_2 = (x_2, y_2, z_2)$ y $\mathbf{r}_3 = (x_3, y_3, z_3)$ de un plano, entonces la ecuación de este plano se representa por el producto triple

$$(\mathbf{r} - \mathbf{r}_1) \cdot (\mathbf{r} - \mathbf{r}_2) \times (\mathbf{r} - \mathbf{r}_3) = 0 \tag{4.45}$$

Sean $\mathbf{a}$ y $\mathbf{b}$ dos vectores cualesquiera no colineales en un plano y sea $\mathbf{r}_1 = (x_1, y_1, z_1)$ el vector de posición de un punto fijo de dicho plano. Los dos vectores

$\mathbf{a}$ y $\mathbf{b}$ pueden considerarse como la base de un sistema de coordenadas $u$ y $v$ en este plano. Entonces,

$$\mathbf{r} - \mathbf{r}_1 = u\mathbf{a} + v\mathbf{b} \quad \text{o} \quad \mathbf{r} = \mathbf{r}_1 + u\mathbf{a} + v\mathbf{b} \tag{4.46}$$

y ésta es la representación paramétrica de un plano.

**Recta**

La intersección de dos planos,

$$\mathbf{A}_1 \cdot \mathbf{r} + D_1 = 0 \quad \text{y} \quad \mathbf{A}_2 \cdot \mathbf{r} + D_2 = 0 \tag{4.47}$$

es una línea **recta**, si las normales $\mathbf{A}_1 = (A_1, B_1, C_1)$ y $\mathbf{A}_2 = (A_2, B_2, C_2)$ a cada superficie no son colineales, es decir, $\mathbf{A}_1 \times \mathbf{A}_2 \neq 0$. Cada línea recta puede representarse en la forma (4.47). La ecuación (4.47) representa una recta que pasa por el origen, si y sólo si $D_1 = D_2 = 0$.

La línea recta que tiene el vector director $\mathbf{a} = (a_x, a_y, a_z)$ y que pasa por el punto $\mathbf{r}_1 = (x_1, y_1, z_1)$ se describe, también, por una ecuación paramétrica

$$\mathbf{r} - \mathbf{r}_1 = \mathbf{a}t \tag{4.48}$$

donde $t$ es una variable (parámetro) real. Aquí se puede tomar $\mathbf{a} = \mathbf{A}_1 \times \mathbf{A}_2$. En coordenadas cartesianas, la ecuación (4.48) toma la forma

$$x = x_1 + a_x t, \quad y = y_1 + a_y t, \quad z = z_1 + a_z t \tag{4.49}$$

Al excluir el parámetro $t$ de las ecuaciones (4.49), se obtiene la ecuación de una recta en forma simétrica

$$\frac{x - x_1}{a_x} = \frac{y - y_1}{a_y} = \frac{z - z_1}{a_z} \tag{4.50}$$

### 4.2.3. Cuádricas

**Ecuación general de segundo orden**

Las **cuádricas** se describen por una ecuación general de segundo orden:

$$\begin{aligned} & a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz + \\ & 2a_{14}x + 2a_{24}y + 2a_{34}z + a_{44} \\ = \; & 0 \end{aligned} \tag{4.51}$$

o

$$\begin{aligned} & (a_{11}x + a_{12}y + a_{13}z + a_{14})\, x + \\ & (a_{21}x + a_{22}y + a_{23}z + a_{24})\, y + \\ & (a_{31}x + a_{32}y + a_{33}z + a_{34})\, z + \\ & (a_{41}x + a_{42}y + a_{43}z + a_{44}) = 0 \end{aligned} \tag{4.52}$$

con $a_{ij} = a_{ji}$, $i, j = 1, 2, 3, 4$.

La forma vectorial de la ecuación (4.51) es

$$\mathbf{r}^T \tilde{A}\mathbf{r} + 2\mathbf{a}\cdot\mathbf{r} + a_{44} = 0$$

donde la matriz $\tilde{A}$ tiene componentes $\tilde{A} = [a_{ik}]$, $i, k = 1, 2, 3$ y el vector $\mathbf{a}$ tiene componentes $a_i = a_{i4}$.

La ecuación (4.51) tiene cuatro invariantes

$$\begin{aligned} I &= a_{11} + a_{22} + a_{33}, \\ J &= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{33} & a_{31} \\ a_{13} & a_{11} \end{vmatrix}, \\ D &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \quad A = \begin{vmatrix} a_{11} & \cdots & a_{14} \\ \vdots & \ddots & \vdots \\ a_{41} & \cdots & a_{44} \end{vmatrix} \end{aligned} \tag{4.53}$$

con respecto a las transformaciones de traslación y rotación. Además, los signos de las magnitudes

$$\begin{aligned} A' &= A_{11} + A_{22} + A_{33} + A_{44}, \\ A'' &= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{14} \\ a_{41} & a_{44} \end{vmatrix} + \\ & \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{22} & a_{24} \\ a_{42} & a_{44} \end{vmatrix} + \begin{vmatrix} a_{33} & a_{34} \\ a_{43} & a_{44} \end{vmatrix}, \\ A''' &= a_{11} + a_{22} + a_{33} + a_{44} \end{aligned} \tag{4.54}$$

son, también, invariantes. Aquí $A_{ij}$ es cofactor de $a_{ij}$ en la matriz $\tilde{A}$ de orden cuatro. Dichos invariantes describen propiedades de la cuádrica que no dependen de su posición en el espacio y, por tanto, permiten la clasificación de las cuádricas en términos de los invariantes. Omitiendo algunos detalles de esta clasificación, nos restringimos en esta sección a las cuádricas propias reales.

Al intentar la cancelación de los términos lineales en la ecuación (4.51) mediante una transformación de traslación del origen del sistema de coordenadas a un punto $(x_0, y_0, z_0)$

$$x = \tilde{x} + x_0, \quad y = \tilde{y} + y_0, \quad z = \tilde{z} + z_0 \tag{4.55}$$

nos encontramos con la necesidad de resolver las ecuaciones

$$\begin{aligned} a_{11}x_0 + a_{12}y_0 + a_{13}z_0 + a_{14} &= 0 \\ a_{21}x_0 + a_{22}y_0 + a_{23}z_0 + a_{24} &= 0 \\ a_{31}x_0 + a_{32}y_0 + a_{33}z_0 + a_{34} &= 0 \end{aligned} \tag{4.56}$$

Si $D \neq 0$, existe la solución única,

$$x_0 = \frac{1}{D}\begin{vmatrix} a_{14} & a_{12} & a_{13} \\ a_{24} & a_{22} & a_{23} \\ a_{34} & a_{32} & a_{33} \end{vmatrix} \tag{4.57}$$

$$y_0 = -\frac{1}{D}\begin{vmatrix} a_{11} & a_{14} & a_{13} \\ a_{21} & a_{24} & a_{23} \\ a_{31} & a_{34} & a_{33} \end{vmatrix} \tag{4.58}$$

$$z_0 = -\frac{1}{D}\begin{vmatrix} a_{11} & a_{12} & a_{14} \\ a_{21} & a_{22} & a_{24} \\ a_{31} & a_{32} & a_{34} \end{vmatrix} \tag{4.59}$$

Las cuádricas con $D \neq 0$ se llaman centrales y el punto $(x_0, y_0, z_0)$, ecuación (4.57), es el centro de la cuádrica. Cuando $D = 0$, las cuádricas no tienen centro y se llaman no centrales. Una cuádrica central en el nuevo sistema de coordenadas toma la forma

$$a_{11}\tilde{x}^2 + a_{22}\tilde{y}^2 + a_{33}\tilde{z}^2 + 2a_{12}\tilde{x}\tilde{y} + 2a_{13}\tilde{x}\tilde{z} + 2a_{23}\tilde{y}\tilde{z} + A/D = 0 \tag{4.60}$$

Propiedades importantes de las cuádricas se pueden estudiar en términos de la **forma cuadrática característica**

$$\begin{aligned} F_0(x,y,z) &= a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz \quad (4.61)\\ F_0(x,y,z) &= \mathbf{r}^T\tilde{A}\mathbf{r} \end{aligned}$$

correspondiente a la ecuación (4.51). En particular, las cuádricas propias centrales ($A \neq 0$, $D \neq 0$) son una elipsoide real o una imaginaria e hiperboloide, si $F_0(x,y,z)$ es positiva, negativa o indefinida, que se determina por las tres raíces $\lambda_1$, $\lambda_2$, $\lambda_3$, necesariamente reales, de la ecuación característica

$$\det\left(\tilde{A} - \lambda I\right) = 0 \quad \Rightarrow \quad \lambda^3 - I\lambda^2 + J\lambda - D = 0 \tag{4.62}$$

En el caso de cuádricas centrales, $D \neq 0$, la ecuación característica tiene tres raíces reales, $\lambda_1$, $\lambda_2$, $\lambda_3$, que determinan tres eigenvectores de la matriz $\tilde{A}$. La rotación del sistema de coordenadas de tal manera que los nuevos ejes estén dirigidos a lo largo de los eigenvectores o sean perpendiculares a ellos, reduce la forma cuadrática característica a la suma de los cuadrados de coordenadas nuevas. Los eigenvectores de la matriz $\tilde{A}$, $\tilde{A}\mathbf{n} = \lambda\mathbf{n}$, son vectores directores de las líneas rectas llamadas **ejes principales** de la **cuádrica**. El plano que pasa por el centro de una cuádrica y tiene como vector normal uno de los eigenvectores de $\tilde{A}$ se llama **plano principal**. Los planos principales son planos de simetría de la cuádrica. Cada cuádrica central tiene cuando menos tres (o un número infinito de) planos principales mutuamente perpendiculares. Las líneas de intersección de los planos principales son los ejes principales, que son ejes de simetría. Cada cuádrica central tiene cuando menos tres (o un número infinito de) ejes principales.

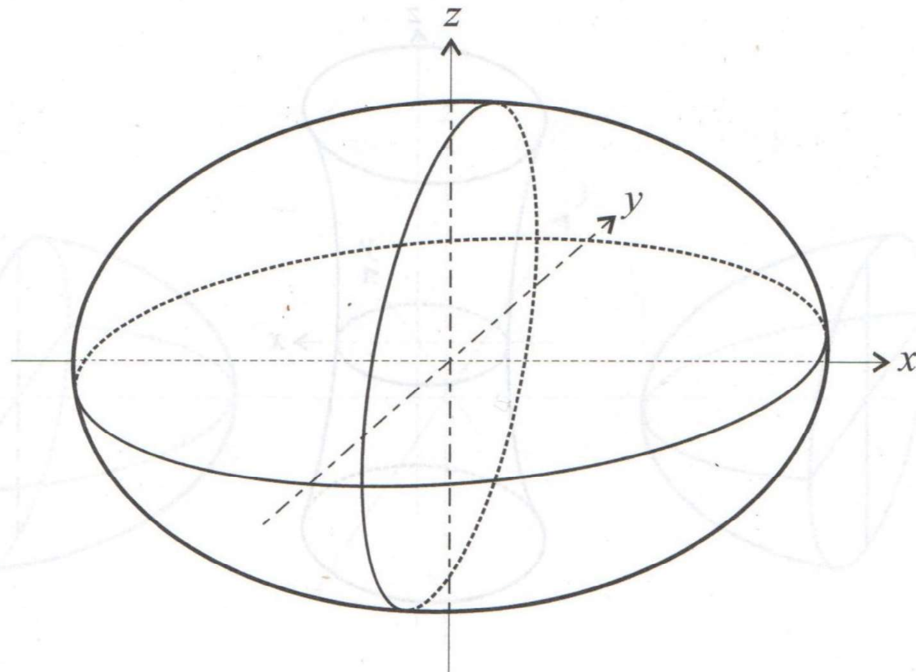



Figura 4.9. Elipsoide

En el caso de las cuádricas no centrales, $D = 0$, la ecuación característica de una cuádrica tiene dos raíces reales no iguales a cero. Además, no existe el centro, no existe la solución de las ecuaciones (4.56), y la ecuación de cuádrica, ecuación (4.51), tiene términos lineales en cualquier sistema de coordenadas. En consecuencia, hay dos (o un número infinito de) planos principales y un solo eje principal.

### Formas canónicas

Mediante una combinación apropiada de una traslación del origen y de una rotación del sistema de coordenadas (**transformación a los ejes principales**) se puede reducir la ecuación (4.51) a una de las formas estándares (**canónicas**); las formas canónicas más frecuentes en la práctica se listan a continuación.

Una **elipsoide**,

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \tag{4.63}$$

es una superficie cuádrica central, en el caso $a \neq b \neq c$ tiene tres planos y tres ejes principales (véase la figura 4.9). La intersección de una elipsoide con cada uno de los planos principales es una elipse. En el caso de $a = b$ una elipsoide se reduce a una superficie de revolución alrededor del eje $z$, y ésta ya tiene un número infinito de planos de simetría.

La **hiperboloide de una capa**,

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 \tag{4.64}$$

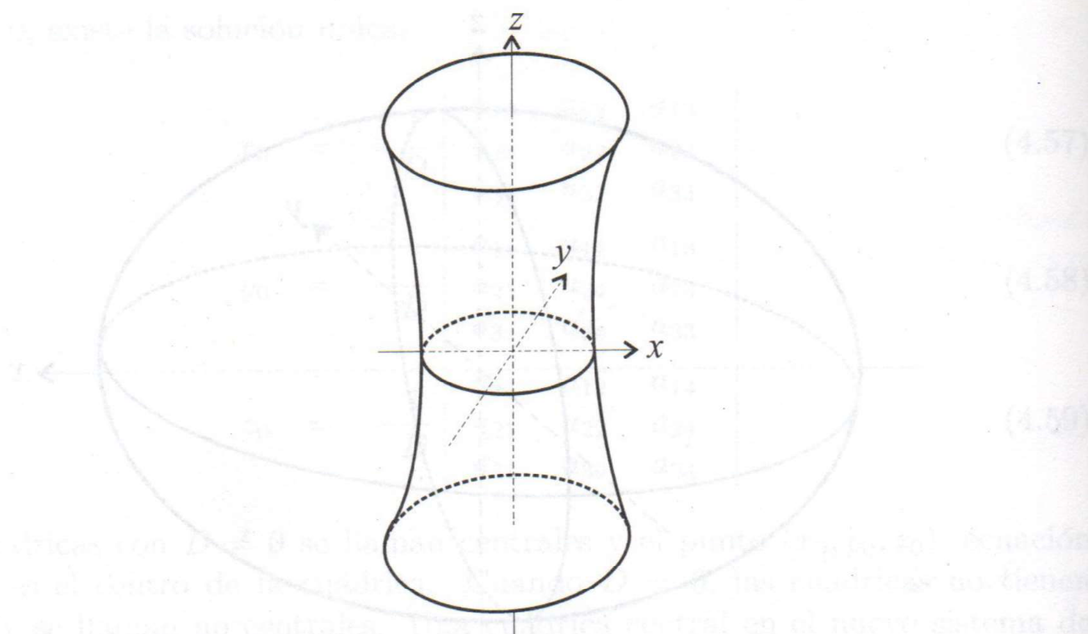



Figura 4.10. Hiperboloide de una capa

es también una cuádrica central, que en el caso de $a \neq b \neq c$ tiene tres planos y tres ejes principales (véase la figura 4.10). La intersección de una hiperboloide de una capa con el plano principal $x = 0$ y el $y = 0$ es una hipérbola; la intersección con el plano $z = 0$ es una elipse. En el caso de $a = b$ una hiperboloide de una capa se reduce a una superficie de revolución alrededor del eje $z$, y ésta ya tiene un número infinito de planos de simetría.

La **hiperboloide de dos capas**,

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 \tag{4.65}$$

es una cuádrica central de dos hojas; en el caso de $a \neq b \neq c$ tiene tres planos y tres ejes principales (véase la figura 4.11). La intersección de una hiperboloide de dos capas con el plano principal $z = 0$ y $y = 0$ es una hipérbola. No existe la intersección real con el plano $x = 0$. En el caso de $b = c$ una hiperboloide de dos capas se reduce a una superficie de revolución alrededor del eje $x$, y ésta ya tiene un número infinito de planos de simetría.

Una **paraboloide elíptica**,

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = z \tag{4.66}$$

es una cuádrica no central; en el caso de $a \neq b$ tiene dos planos y un solo eje principal (véase la figura 4.12). La intersección de una paraboloide elíptica con el plano principal $x = 0$ y $y = 0$ es una parábola. La intersección con el plano $z = const > 0$ es una elipse, no existe la intersección real con el plano $z = const < 0$. En el caso de $a = b$ una paraboloide elíptica se reduce a una superficie de revolución alrededor del eje $z$, y ésta ya tiene un número infinito de planos de simetría y un solo eje principal.

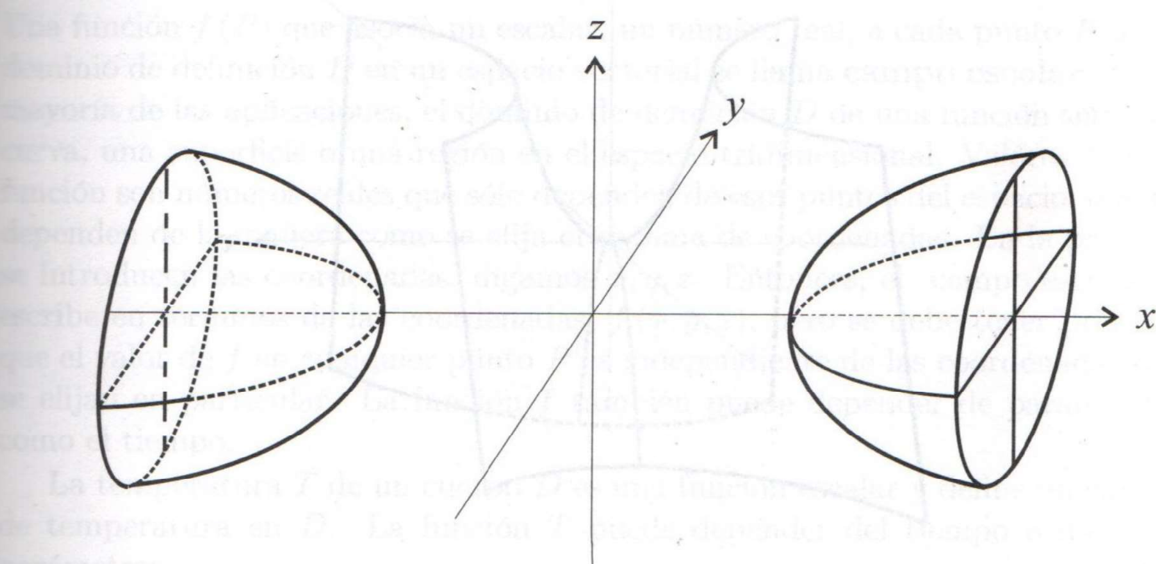



Figura 4.11. Hiperboloide de dos capas

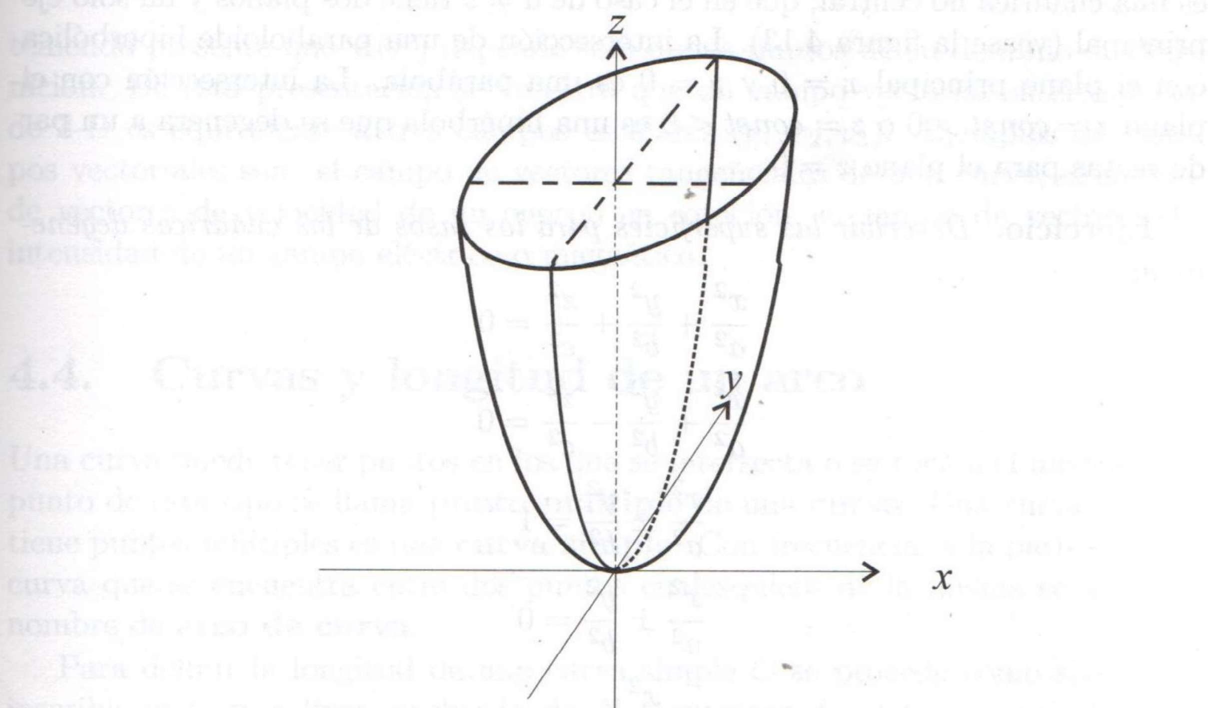



Figura 4.12. Paraboloide elíptico

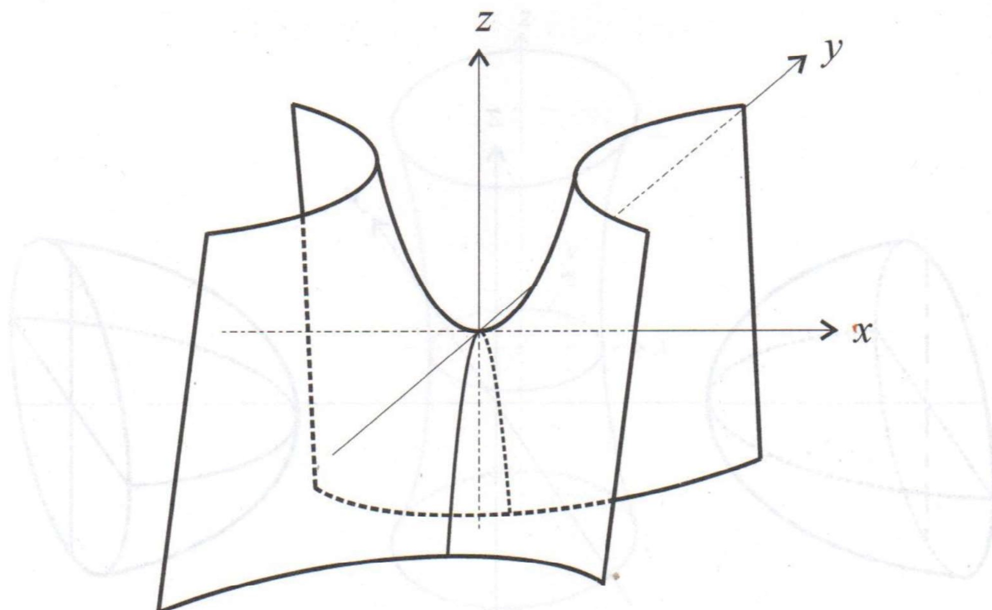



Figura 4.13. Paraboloide hiperbólico

**Paraboloide hiperbólica** o la silla del jinete,

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = z \tag{4.67}$$

es una cuádrica no central, que en el caso de $a \neq b$ tiene dos planos y un solo eje principal (véase la figura 4.13). La intersección de una paraboloide hiperbólica con el plano principal $x = 0$ y $y = 0$ es una parábola. La intersección con el plano $z = const > 0$ o $z = const < 0$ es una hipérbola que se degenera a un par de rectas para el plano $z = 0$.

**Ejercicio.** *Describir las superficies para los casos de las cuádricas degeneradas:*

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 0$$

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0$$

$$\frac{x^2}{a^2} \pm \frac{y^2}{b^2} = 1$$

$$\frac{x^2}{a^2} \pm \frac{y^2}{b^2} = 0$$

$$\frac{x^2}{a^2} = y$$

$$x^2 = a^2$$

$$x^2 = 0$$

## 4.3. Campos escalares y vectoriales

Una función $f(P)$ que asocia un escalar, un número real, a cada punto $P$ de su dominio de definición $D$ en un espacio vectorial se llama **campo escalar**. En la mayoría de las aplicaciones, el dominio de definición $D$ de una función será una curva, una superficie o una región en el espacio tridimensional. Valores de esta función son números reales que sólo dependen de esos puntos del espacio, que no dependen de la manera como se elija el sistema de coordenadas. En la práctica se introducen las coordenadas, digamos $x, y, z$. Entonces, el campo escalar se escribe en términos de las coordenadas, $f(x, y, z)$, pero se debe tener presente que el valor de $f$ en cualquier punto $P$ es independiente de las coordenadas que se elijan en particular. La función $f$ también puede depender de parámetros, como el tiempo.

La temperatura $T$ de un cuerpo $D$ es una función escalar y define un campo de temperatura en $D$. La función $T$ puede depender del tiempo o de otros parámetros.

Si a cada punto $P$ de un dominio $D$ en el espacio se le asigna un vector $\mathbf{v}(P)$, se dice que está definida una **función vectorial** en el dominio $D$. Los valores vectoriales de $\mathbf{v}(P)$ forman cierto conjunto $\tilde{D}$ que se llama **campo vectorial** en los puntos de $D$. Si se introducen las coordenadas cartesianas $x, y, z$, los vectores $\mathbf{v}(P)$ pueden presentarse en la forma

$$\mathbf{v}(P) = \mathbf{v}(x, y, z) = v_1(x, y, z)\,\mathbf{i} + v_2(x, y, z)\,\mathbf{j} + v_3(x, y, z)\,\mathbf{k} \tag{4.68}$$

teniendo presente que $\mathbf{v}(P)$ depende sólo de los puntos de su dominio de definición. De esta presentación se ve claro que un campo vectorial en el espacio de 3-D es equivalente a tres campos escalares $(v_1, v_2, v_3)$. Ejemplos de campos vectoriales son: el campo de vectores tangenciales de una curva, el campo de vectores de velocidad de un cuerpo en rotación, el campo de vectores de intensidad de un campo eléctrico o magnético.

## 4.4. Curvas y longitud de un arco

Una curva puede tener puntos en los que se intersecta o se toca a sí misma. Un punto de este tipo se llama **punto múltiple** de una **curva**. Una curva que no tiene puntos múltiples es una **curva simple**. Con frecuencia, a la parte de una curva que se encuentra entre dos puntos cualesquiera de la misma se le da el nombre de **arco de curva**.

Para definir la longitud de una curva simple $C$ se procede como sigue. Se inscribe en $C$ una línea quebrada de $N$ segmentos de recta que una los dos puntos extremos de $C$. Después, el número $N$ de quebraduras se hace crecer de tal manera que la longitud del segmento más grande de recta tienda a cero conforme $N \to \infty$. La longitud de cada línea quebrada es la suma de las longitudes de los segmentos de recta que constituyen la quebrada, y la longitud de cada segmento de recta se define fácil por el teorema de Pitágoras. Si la

sucesión de longitudes de las líneas quebradas converge a un límite $L$, se dice que una **curva** $C$ es **rectificable** y la longitud de $C$ es igual a $L$.

Si una curva $C$ no es simple pero consta de un número finito de curvas simples rectificables, la longitud de $C$ se define como la suma de las longitudes de estas curvas.

Suponemos que una curva $C$ está representada mediante una función vectorial continuamente diferenciable,

$$\mathbf{r} = \mathbf{r}(t) = x(t)\,\mathbf{i} + y(t)\,\mathbf{j} + z(t)\,\mathbf{k} \tag{4.69}$$

de un parámetro $a \leq t \leq b$. Dividimos el intervalo $a = t_0 < t_1 < \cdots < t_N = b$ en $N$ subintervalos y unimos los pares de puntos consecutivos de $C$ con segmentos de recta, obteniendo de esta manera una línea quebrada inscrita en $C$. El segmento $k$-ésimo $d\mathbf{r}_k$ de la línea quebrada inscrita en $C$ es

$$\begin{aligned} d\mathbf{r}_k &= \mathbf{r}(t_k) - \mathbf{r}(t_{k-1}) \simeq dx(t_{k-1})\,\mathbf{i} + dy(t_{k-1})\,\mathbf{j} + dz(t_{k-1})\,\mathbf{k} \\ &= \left[\dot{x}(t_{k-1})\,\mathbf{i} + \dot{y}(t_{k-1})\,\mathbf{j} + \dot{z}(t_{k-1})\,\mathbf{k}\right]\Delta t_k \end{aligned} \tag{4.70}$$

donde $\Delta t_k = t_k - t_{k-1}$, $k = 1, \ldots, N$. El vector $d\mathbf{r}_k$ tiene la longitud

$$ds_k = (d\mathbf{r}_k \cdot d\mathbf{r}_k)^{1/2} \simeq \left[(\dot{x}(t_{k-1}))^2 + (\dot{y}(t_{k-1}))^2 + \left(\dot{z}(t_{k-1})^2\right)\right]^{1/2} \Delta t_k \tag{4.71}$$

La longitud $L_N$ de la quebrada de $N$ segmentos es la suma de las longitudes de dichos segmentos. La longitud de $C$ es el límite de $L_N$ cuando $N \to \infty$, es decir,

$$\begin{aligned} L &= \lim_{N\to\infty} L_N = \lim_{N\to\infty} \sum_{k=1}^{N} ds_k \\ &= \int_a^b \sqrt{\dot{\mathbf{r}} \cdot \dot{\mathbf{r}}}dt = \int_a^b \left[(\dot{x}(t))^2 + (\dot{y}(t))^2 + \left(\dot{z}(t)^2\right)\right]^{1/2} dt \end{aligned} \tag{4.72}$$

Reemplazando el límite superior fijo $b$ en la ecuación (4.72) por un límite superior variable $t$, se obtiene la integral como función de $t$,

$$s(t) = \int_a^t \sqrt{\dot{\mathbf{r}} \cdot \dot{\mathbf{r}}}d\tau \tag{4.73}$$

Esta función es la **longitud de arco** de $C$. El uso de $s$ como parámetro natural en la representación de una curva $C$ simplifica diversas fórmulas. El sentido que corresponde a los valores crecientes de $s$ se llama sentido positivo sobre $C$; de esta manera se define cierta orientación de la curva.

Un cambio de parametrización de $C$, digamos $s = s(t)$, reemplazará la ecuación (4.72) por la ecuación

$$L = \int_{s(a)}^{s(b)} \sqrt{\dot{\mathbf{r}}(s) \cdot \dot{\mathbf{r}}(s)}ds \tag{4.74}$$

Si $s$ es el parámetro natural de $C$, $\mathbf{r} = \mathbf{r}(s)$, entonces, $s(a) = 0$ y se tiene la identidad

$$s = \int_0^s \sqrt{\dot{\mathbf{r}}(\tilde{s}) \cdot \dot{\mathbf{r}}(\tilde{s})} d\tilde{s} \tag{4.75}$$

denotando la variable de integración por $\tilde{s}$. De la ecuación (4.73) se obtiene

$$\left(\frac{ds}{dt}\right)^2 = \frac{d\mathbf{r}}{dt} \cdot \frac{d\mathbf{r}}{dt} = \left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2 \tag{4.76}$$

o en diferenciales,

$$(ds)^2 = d\mathbf{r} \cdot d\mathbf{r} = (dx)^2 + (dy)^2 + (dz)^2 \tag{4.77}$$

# 4.5. Gradiente de un campo escalar

### Definición

Sea $\varphi(x, y, z)$ una función escalar (campo escalar) diferenciable definida en el espacio de 3-D, en donde $x, y, z$ son coordenadas cartesianas de un punto en dicho espacio. Del Cálculo diferencial se conoce que un incremento infinitesimal de $\varphi$ a causa de pequeños incrementos de sus argumentos está dado por la fórmula

$$d\varphi = \frac{\partial \varphi}{\partial x} dx + \frac{\partial \varphi}{\partial y} dy + \frac{\partial \varphi}{\partial z} dz \tag{4.78}$$

Sea $\mathbf{r}$ el vector de posición de un punto $P(x, y, z)$, $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$. El desplazamiento a otro punto $Q(x + dx, y + dy, z + dz)$ se representa por el vector $d\mathbf{r} = dx\mathbf{i} + dy\mathbf{j} + dz\mathbf{k}$. Nótese que la ecuación (4.78) contiene tanto los términos $dx, dy, dz$ como los términos $\partial\varphi/\partial x, \partial\varphi/\partial y, \partial\varphi/zx$. Definimos un nuevo vector, llamado el **gradiente** de $\varphi$, mediante la fórmula

$$\text{grad}\varphi \equiv \boldsymbol{\nabla}\varphi = \frac{\partial \varphi}{\partial x}\mathbf{i} + \frac{\partial \varphi}{\partial y}\mathbf{j} + \frac{\partial \varphi}{\partial z}\mathbf{k} \tag{4.79}$$

De inmediato se ve que

$$d\varphi = d\mathbf{r} \cdot \boldsymbol{\nabla}\varphi \tag{4.80}$$

### Interpretación geométrica del gradiente

En el punto $P(x_0, y_0, z_0)$ la función $\varphi$ tiene el valor $\varphi(x_0, y_0, z_0)$. La ecuación $\varphi(x, y, z) = \varphi(x_0, y_0, z_0)$ representa una superficie que contiene a $P(x_0, y_0, z_0)$. Mientras nos desplazamos sobre esta superficie, la función $\varphi$ tiene el mismo valor constante $\varphi(x_0, y_0, z_0)$ y, por lo tanto, $d\varphi = 0$. De la ecuación (4.80) se tiene

$$d\mathbf{r} \cdot \boldsymbol{\nabla}\varphi = 0 \tag{4.81}$$

donde el desplazamiento $dr$ es sobre dicha superficie. El gradiente $\boldsymbol{\nabla}\varphi$ es completamente determinado por la función $\varphi$. Entonces, la ecuación (4.81) establece

que $\nabla\varphi$ es perpendicular a $d\mathbf{r}$ mientras que el vector $d\mathbf{r}$ une dos puntos $P$ y $Q$ que pertenecen a la superficie $\varphi(x,y,z) = const$. Entonces, $\nabla\varphi$ es perpendicular a todas las tangentes a dicha superficie en el punto $P$, es decir, $\nabla\varphi$ es una normal a la superficie $\varphi(x,y,z) = const$.

Para un punto dado $P(x,y,z)$, el vector $\nabla\varphi$ es fijo. Por tanto, $d\varphi$ dependerá solamente de $d\mathbf{r}$. Sea $ds = |d\mathbf{r}|$, pues $d\mathbf{r} = \mathbf{u}\cdot ds$, donde $\mathbf{u}$ es el vector unitario en la dirección de $d\mathbf{r}$. Luego,

$$d\varphi = ds\,(\mathbf{u}\cdot\nabla\varphi) \quad \text{y} \quad \frac{d\varphi}{ds} = \mathbf{u}\cdot\nabla\varphi \tag{4.82}$$

La derivada $d\varphi/ds$ se llama **derivada direccional** en la dirección de $\mathbf{u}$. Si la dirección de $\mathbf{u}$ coincide con la del gradiente $\nabla\varphi$, entonces, la derivada direccional alcanza su máximo valor. Así, el gradiente nos indica la dirección del máximo cambio de la función $\varphi$.

**Ejemplo 4.1.** *Sea $\varphi(x,y,z) = r = |\mathbf{r}| = \sqrt{x^2+y^2+z^2}$. Por lo tanto, la superficie $\varphi(x,y,z) = const$ es una esfera. Dado que $\nabla\varphi$ es una normal a la superficie $\varphi(x,y,z) = const$, tenemos la igualdad*

$$\nabla r = k\mathbf{r}$$

*donde $k$ es un escalar por determinar. Ya que $r^2 = \mathbf{r}\cdot\mathbf{r}$, se tiene que $d(r^2) = 2rdr = 2\mathbf{r}\cdot d\mathbf{r}$. Así que*

$$d\varphi = dr = d\mathbf{r}\cdot\nabla r = kd\mathbf{r}\cdot\mathbf{r} = krdr$$

*de donde $k = 1/r$. Finalmente,*

$$\nabla r = \frac{\mathbf{r}}{r} \tag{4.83}$$

*es un vector unitario.*

**Ejemplo 4.2.** *Demostraremos que*

$$\nabla f(u) = \frac{df(u)}{du}\nabla u, \quad \text{donde} \quad u = u(x,y,z) \tag{4.84}$$

*Por definición,*

$$\begin{aligned}
\nabla f(u) &= \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k} \\
&= \frac{df}{du}\frac{\partial u}{\partial x}\mathbf{i} + \frac{df}{du}\frac{\partial u}{\partial y}\mathbf{j} + \frac{df}{du}\frac{\partial u}{\partial z}\mathbf{k} \\
&= \frac{df}{du}\nabla u
\end{aligned}$$

**Operador diferencial $\nabla$**

Definimos el operador diferencial vectorial

$$\boldsymbol{\nabla} = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial x} + \mathbf{k}\frac{\partial}{\partial x} \tag{4.85}$$

de tal manera que

$$\boldsymbol{\nabla}\varphi = \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial x} + \mathbf{k}\frac{\partial}{\partial x}\right)\varphi = \mathbf{i}\frac{\partial\varphi}{\partial x} + \mathbf{j}\frac{\partial\varphi}{\partial x} + \mathbf{k}\frac{\partial\varphi}{\partial x}$$

El **operador** $\boldsymbol{\nabla}$, llamado **nabla**, tiene las características tanto de un vector como de un operador diferencial. Por lo tanto,

$$\begin{aligned}\boldsymbol{\nabla}(uv) &= \mathbf{i}\frac{\partial(uv)}{\partial x} + \mathbf{j}\frac{\partial(uv)}{\partial x} + \mathbf{k}\frac{\partial(uv)}{\partial x} \\ &= u\boldsymbol{\nabla}v + v\boldsymbol{\nabla}u\end{aligned} \tag{4.86}$$

y

$$\boldsymbol{\nabla}(\mathbf{a}\cdot\mathbf{r}) = \mathbf{a} \tag{4.87}$$

donde $\mathbf{a} = (a_x, a_y, a_z)$ es un vector constante.

## 4.6. Coordenadas curvilíneas ortogonales

A menudo los matemáticos, físicos o ingenieros se encuentran ante el problema de usar un sistema de coordenadas diferente del sistema cartesiano rectangular. La razón principal es que por la naturaleza de un problema físico o matemático a resolver, tanto el planteamiento como la resolución del problema resultan más simples en un sistema "natural" para el problema.

### 4.6.1. Coordenadas esféricas

Un problema físico con la simetría esférica es mucho más fácil de resolver en el sistema de coordenadas esféricas, que es natural para el problema. En este caso, la posición de un punto en el espacio físico se describe mediante las coordenadas esféricas $r, \theta, \varphi$ que se definen mediante el siguiente procedimiento (véase la figura 4.14). Notamos que la esfera $x^2 + y^2 + z^2 = r^2$, donde $r = |\mathbf{r}|$ es el módulo del vector de posición, el cono $z/(x^2 + y^2 + z^2)^{1/2} = \cos\theta$ y el plano $y/x = \tan\varphi$ pasan por el punto $P(r, \theta, \varphi)$. Por lo tanto, las relaciones

$$r = \sqrt{x^2 + y^2 + z^2} \tag{4.88}$$

$$\theta = \arccos\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right) \tag{4.89}$$

$$\varphi = \arctan\frac{y}{x} \tag{4.90}$$

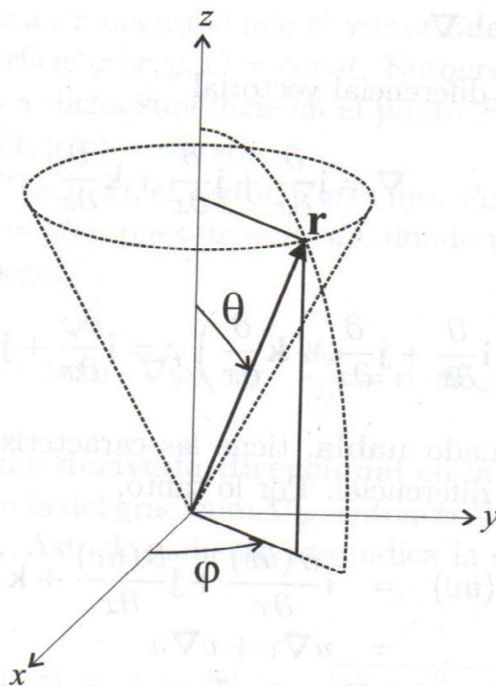



Figura 4.14. Coordenadas esféricas

pueden considerarse como la transformación del sistema de coordenadas cartesianas $x, y, z$ al sistema de coordenadas esféricas $r, \theta, \varphi$. A través de cada punto del espacio, salvo el origen, pasará una superficie del tipo $r = \sqrt{x^2 + y^2 + z^2} = c_1$, $\theta = \arccos(z/r) = c_2$ y $\varphi = \arctan(y/x) = c_3$, que son una esfera, un cono y un plano, respectivamente. Las coordenadas del punto $P$ se determinarán por las constantes $c_1, c_2, c_3$.

La intersección de la esfera con el cono es una circunferencia de radio $r \operatorname{sen} \theta$ a una altura $z = r\cos\theta$ que tiene un vector tangente unitario $\mathbf{e}_\varphi$ en el punto $P$. Esta circunferencia se llama curva-$\varphi$ ($r = c_1$, $\theta = c_2$) porque $r$ y $\theta$ se mantienen fijos sobre ésta y sólo la coordenada $\varphi$ se cambia al moverse a lo largo de la curva. La intersección de la esfera y el plano que pasa por el origen resulta en una circunferencia llamada curva-$\theta$ ($r = c_1$, $\varphi = c_3$), mientras que la intersección del cono con el plano resulta en la línea recta, curva-$r$ ($\theta = c_2$, $\varphi = c_3$), que pasa por el origen y el punto $P$. Sean $\mathbf{e}_\theta$ y $\mathbf{e}_r$ vectores tangenciales unitarios en el punto $P$ a las curvas -$\theta$ y -$r$, respectivamente. Es fácil ver que las tres superficies $r = c_1$, $\theta = c_2$ y $\varphi = c_3$ son perpendiculares entre sí en cada punto de intersección. Entonces, los tres vectores $\mathbf{e}_r$, $\mathbf{e}_\theta$, $\mathbf{e}_\varphi$ son perpendiculares entre sí y, por tanto, forman una base ortonormal en la vecindad del punto $P$. Dicha base es dependiente de cada punto $P$ y, en este sentido, el sistema de coordenadas es local.

Dado que tanto $\mathbf{e}_\varphi$ como $\nabla\varphi$ son perpendiculares a la superficie $\varphi = const$ en el punto $P$, el vector $\nabla\varphi$ debe ser paralelo a $\mathbf{e}_\varphi$. Entonces, $\mathbf{e}_\varphi = h_3 \nabla\varphi$, en donde $h_3$ es el factor escalar de proporcionalidad entre $\mathbf{e}_\varphi$ y $\nabla\varphi$. Si $d\mathbf{r}_3$ es un vector infinitesimal tangente a la curva-$\varphi$ de longitud $ds_3 = |d\mathbf{r}_3|$, entonces se tiene

$$ds_3 = d\mathbf{r}_3 \cdot \mathbf{e}_\varphi = d\mathbf{r}_3 \cdot (h_3 \boldsymbol{\nabla}\varphi) = h_3 d\varphi$$

ya que por la fórmula (4.80) $d\varphi = d\mathbf{r}_3 \cdot \boldsymbol{\nabla}\varphi$. Se ve que $h_3$ representa el factor que, al haber multiplicado por el cambio diferencial $d\varphi$ de la coordenada $\varphi$, nos da la longitud del arco a lo largo de la curva-$\varphi$. Entonces, $h_3 = r \operatorname{sen}\theta$ y

$$\mathbf{e}_\varphi = h_3 \boldsymbol{\nabla}\varphi = r \operatorname{sen}\theta \boldsymbol{\nabla}\varphi \tag{4.91}$$

De manera similar se comprueba que

$$\mathbf{e}_\theta = h_2 \boldsymbol{\nabla}\theta = r\boldsymbol{\nabla}\theta \tag{4.92}$$

$$\mathbf{e}_r = h_1 \boldsymbol{\nabla} r = \boldsymbol{\nabla} r \tag{4.93}$$

Notamos que

$$\begin{aligned}
\mathbf{e}_r &= \mathbf{e}_\theta \times \mathbf{e}_\varphi = r^2 \operatorname{sen}\theta \,(\boldsymbol{\nabla}\theta \times \boldsymbol{\nabla}\varphi) \\
\mathbf{e}_\theta &= \mathbf{e}_\varphi \times \mathbf{e}_r = r \operatorname{sen}\theta \,(\boldsymbol{\nabla}\theta \times \boldsymbol{\nabla} r) \\
\mathbf{e}_\varphi &= \mathbf{e}_r \times \mathbf{e}_\theta = r \,(\boldsymbol{\nabla} r \times \boldsymbol{\nabla}\varphi)
\end{aligned}$$

Cada vector $\mathbf{f}$ en el punto $P$ puede representarse en la base local como

$$\mathbf{f} = f_r \mathbf{e}_r + f_\theta \mathbf{e}_\theta + f_\varphi \mathbf{e}_\varphi = f_r \boldsymbol{\nabla} r + r f_\theta \boldsymbol{\nabla}\theta + f_\varphi r \operatorname{sen}\theta \boldsymbol{\nabla}\varphi$$

en donde las coordenadas $f_r, f_\theta, f_\varphi$ pueden ser funciones de $r, \theta, \varphi$. Además, la diferencial de volumen es como sigue:

$$\begin{aligned}
dV &= dxdydz = \Im(x, y, z/r, \theta, \varphi)\, drd\theta d\varphi \\
&= d\mathbf{r}_1 \cdot d\mathbf{r}_2 \times d\mathbf{r}_3 = ds_1 ds_2 ds_3 \mathbf{e}_r \cdot \mathbf{e}_\theta \times \mathbf{e}_\varphi = ds_1 ds_2 ds_3 \\
&= h_1 dr h_2 d\theta h_3 d\varphi = r^2 \operatorname{sen}\theta drd\theta d\varphi
\end{aligned} \tag{4.94}$$

en donde $\Im(x, y, z/r, \theta, \varphi)$ es el Jacobiano de la transformada $x, y, z \to r, \theta, \varphi$.

### 4.6.2. Coordenadas curvilíneas ortogonales: conceptos generales

Las coordenadas esféricas son un caso particular de los sistemas de coordenadas curvilíneas ortogonales. En esta sección se consideran sistemas de coordenadas curvilíneas ortogonales en general, con el fin de tener las representaciones del gradiente, la divergencia, el rotacional y el laplaciano en cualquiera de estos sistemas.

Sea

$$u_1 = u_1(x, y, z), \quad u_2 = u_2(x, y, z), \quad u_3 = u_3(x, y, z) \tag{4.95}$$

un cambio del sistema de coordenadas cartesiano $x, y, z$ a un sistema $u_1, u_2, u_3$. Suponemos que el Jacobiano $\Im(u_1, u_2, u_3/x, y, z) \neq 0$ en un dominio $D$ de espacio y, por lo tanto, la transformación de coordenadas (4.95) es uno a uno en $D$. Las ecuaciones

$$u_1(x, y, z) = c_1, \quad u_2(x, y, z) = c_2, \quad u_3(x, y, z) = c_3 \tag{4.96}$$

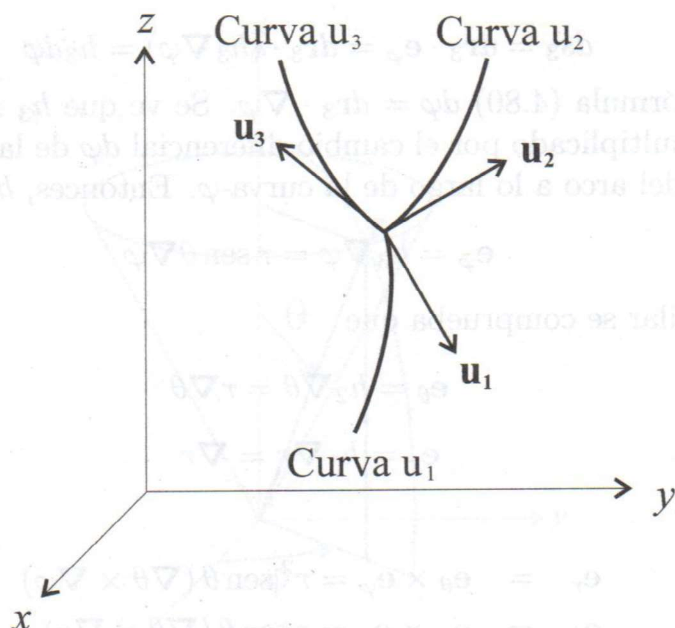



Figura 4.15. Coordenadas curvilíneas ortogonales

en donde $c_1, c_2, c_3$ son constantes, determinan una familia de superficies. A través de cada punto $P(x_0, y_0, z_0)$ del dominio $D$ pasarán las tres superficies

$$\begin{aligned} u_1(x,y,z) &= u_1(x_0,y_0,z_0) \\ u_2(x,y,z) &= u_2(x_0,y_0,z_0) \\ u_3(x,y,z) &= u_3(x_0,y_0,z_0) \end{aligned}$$

Suponemos que estas superficies son ortogonales entre sí en cada punto $P(x_0, y_0, z_0)$, es decir, las normales de las superficies son ortogonales entre sí en cada punto. En consecuencia, las tres curvas de intersección de cada par de superficies son ortogonales entre sí en cada punto $P(x_0, y_0, z_0)$. La curva de intersección de las superficies $u_1 = c_1$ y $u_2 = c_2$ la llamaremos curva-$u_3$, dado que a lo largo de dicha curva se cambia sólo la variable $u_3$. Respectivamente, la curva-$u_1$ es la intersección de $u_2 = c_2$ y $u_3 = c_3$, la curva-$u_2$ es la intersección de $u_1 = c_1$ y $u_3 = c_3$. Sean $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ los tres vectores unitarios tangentes a las curvas-$u_1$, -$u_2$ y -$u_3$ en el punto $P$. Dado que las superficies $u_1 = c_1$, $u_2 = c_2$ y $u_3 = c_3$ son ortogonales entre sí en cada punto de su intersección, los tres vectores $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ son ortogonales entre sí y, por lo tanto, forman una base ortonormal (véase la figura 4.15).

El vector $\nabla u_1$ es perpendicular a la superficie $u_1(x,y,z) = u_1(x_0,y_0,z_0)$ y, por lo tanto, es paralelo al vector $\mathbf{u}_1$, es decir,

$$\mathbf{u}_1 = h_1 \nabla u_1 \tag{4.97}$$

en donde $h_1$ es el factor escalar de proporcionalidad entre los $\mathbf{u}_1$ y $\nabla u_1$. Sea $d\mathbf{r}_1$ un vector infinitesimal tangente a la curva-$u_1$ de longitud $ds_1 = |d\mathbf{r}_1|$. Es claro que $ds_1 = d\mathbf{r}_1 \cdot \mathbf{u}_1$. Entonces, según (4.80), se tiene

$$ds_1 = d\mathbf{r}_1 \cdot \mathbf{u}_1 = h_1 d\mathbf{r}_1 \cdot \nabla u_1 = h_1 du_1 \tag{4.98}$$

De aquí se comprende que $h_1$ multiplicada por la diferencial de coordenada $du_1$ nos da la longitud del arco. Por ejemplo, en el caso de coordenadas esféricas $ds = rd\theta$ al desplazarse a lo largo de la curva-$\theta$, es decir, $h_2 = r$. De igual forma,

$$\mathbf{u}_2 = h_2 \nabla u_2, \quad \mathbf{u}_3 = h_3 \nabla u_3 \tag{4.99}$$

Luego,

$$\begin{aligned} \mathbf{u}_1 &= \mathbf{u}_2 \times \mathbf{u}_3 = h_2 h_3 \nabla u_2 \times \nabla u_3 \\ \mathbf{u}_2 &= \mathbf{u}_3 \times \mathbf{u}_1 = h_3 h_1 \nabla u_3 \times \nabla u_1 \\ \mathbf{u}_3 &= \mathbf{u}_1 \times \mathbf{u}_2 = h_1 h_2 \nabla u_1 \times \nabla u_2 \end{aligned} \tag{4.100}$$

y

$$\nabla u_1 \left(\nabla u_2 \times \nabla u_3\right) = \frac{\mathbf{u}_1 \cdot \mathbf{u}_2 \times \mathbf{u}_3}{h_1 h_2 h_3} = \frac{1}{h_1 h_2 h_3} \tag{4.101}$$

La diferencial de volumen es

$$\begin{aligned} dV &= \Im\left(x, y, z / u_1, u_2, u_3\right) du_1 du_2 du_3 \\ &= ds_1 ds_2 ds_3 = h_1 h_2 h_3 du_1 du_2 du_3 \end{aligned} \tag{4.102}$$

de donde,

$$\Im\left(x, y, z / u_1, u_2, u_3\right) = h_1 h_2 h_3 \tag{4.103}$$

### Gradiente

Si $f\left(u_1, u_2, u_3\right)$ es un campo escalar, entonces, con base en (4.84) se tiene

$$\begin{aligned} \nabla f &= \frac{\partial f}{\partial u_1} \nabla u_1 + \frac{\partial f}{\partial u_2} \nabla u_2 + \frac{\partial f}{\partial u_3} \nabla u_3 \\ &= \frac{1}{h_1} \frac{\partial f}{\partial u_1} \mathbf{u}_1 + \frac{1}{h_2} \frac{\partial f}{\partial u_2} \mathbf{u}_2 + \frac{1}{h_3} \frac{\partial f}{\partial u_3} \mathbf{u}_3 \end{aligned} \tag{4.104}$$

donde $u_1 = u_1\left(x, y, z\right)$, $u_2 = u_2\left(x, y, z\right)$ y $u_3 = u_3\left(x, y, z\right)$.

### Divergencia

Sea $\mathbf{f}\left(u_1, u_2, u_3\right)$ un campo vectorial cuyos componentes son funciones conocidas $f_1\left(u_1, u_2, u_3\right)$, $f_2\left(u_1, u_2, u_3\right)$ y $f_3\left(u_1, u_2, u_3\right)$ en un sistema ortogonal de coordenadas $u_1, u_2, u_3$. Entonces, de (4.100)

$$\begin{aligned} \mathbf{f}\left(u_1, u_2, u_3\right) &= f_1 \mathbf{u}_1 + f_2 \mathbf{u}_2 + f_3 \mathbf{u}_3 \\ &= f_1 h_2 h_3 \nabla u_2 \times \nabla u_3 + f_2 h_3 h_1 \nabla u_3 \times \nabla u_1 + \\ &\quad f_3 h_1 h_2 \nabla u_1 \times \nabla u_2 \end{aligned}$$

De donde, usando la fórmula

$$\nabla \cdot \left(\varphi \mathbf{u}\right) = \varphi \nabla \cdot \mathbf{u} + \left(\nabla \varphi\right) \cdot \mathbf{u} \tag{4.105}$$

se tiene

$$\begin{aligned}\nabla\cdot\mathbf{f} &= f_1h_2h_3\left(\nabla\cdot(\nabla u_2\times\nabla u_3)\right)+\nabla(f_1h_2h_3)\cdot(\nabla u_2\times\nabla u_3)+\\ &\quad f_2h_3h_1\left(\nabla\cdot(\nabla u_3\times\nabla u_1)\right)+\nabla(f_2h_3h_1)\cdot(\nabla u_3\times\nabla u_1)+\\ &\quad f_3h_1h_2\left(\nabla\cdot(\nabla u_1\times\nabla u_2)\right)+\nabla(f_3h_1h_2)\cdot(\nabla u_1\times\nabla u_2)\end{aligned} \quad (4.106)$$

Luego, aplicando las fórmulas

$$\nabla\cdot(\mathbf{u}\times\mathbf{v})=(\nabla\times\mathbf{u})\cdot\mathbf{v}-(\nabla\times\mathbf{v})\cdot\mathbf{u} \quad (4.107)$$

y

$$\nabla\times(\nabla\varphi)=0 \quad (4.108)$$

de (4.106) se obtiene

$$\begin{aligned}\nabla\cdot\mathbf{f} &= \nabla(f_1h_2h_3)\cdot(\nabla u_2\times\nabla u_3)+\\ &\quad \nabla(f_2h_3h_1)\cdot(\nabla u_3\times\nabla u_1)+\\ &\quad \nabla(f_3h_1h_2)\cdot(\nabla u_1\times\nabla u_2)\end{aligned}$$

De la última, con base en (4.97), (4.99) y (4.104), finalmente se obtiene la divergencia

$$\nabla\cdot\mathbf{f}=\frac{1}{h_1h_2h_3}\left[\frac{\partial(f_1h_2h_3)}{\partial u_1}+\frac{\partial(f_2h_3h_1)}{\partial u_2}+\frac{\partial(f_3h_1h_2)}{\partial u_3}\right] \quad (4.109)$$

**Laplaciano**

Aplicando (4.109) al vector $\nabla V$, en donde $V(u_1,u_2,u_3)$ es un campo escalar, se obtiene

$$\nabla^2V=\frac{1}{h_1h_2h_3}\left[\frac{\partial}{\partial u_1}\left(\frac{h_2h_3}{h_1}\frac{\partial V}{\partial u_1}\right)+\frac{\partial}{\partial u_2}\left(\frac{h_3h_1}{h_2}\frac{\partial V}{\partial u_2}\right)+\frac{\partial}{\partial u_3}\left(\frac{h_1h_2}{h_3}\frac{\partial V}{\partial u_3}\right)\right] \quad (4.110)$$

que es el laplaciano en un sistema ortogonal de coordenadas curvilíneas. El laplaciano en coordenadas esféricas con $h_1=1$, $h_2=r$ y $h_3=r\operatorname{sen}\theta$ es

$$\nabla^2V=\frac{1}{r^2}\left[\frac{\partial}{\partial r}\left(r^2\frac{\partial V}{\partial r}\right)+\frac{1}{\operatorname{sen}\theta}\frac{\partial}{\partial\theta}\left(\operatorname{sen}\theta\frac{\partial V}{\partial\theta}\right)+\frac{1}{\operatorname{sen}^2\theta}\frac{\partial}{\partial\varphi}\left(\frac{\partial V}{\partial\varphi}\right)\right] \quad (4.111)$$

## 4.6.3. Coordenadas cilíndricas

Las coordenadas cilíndricas $u_1\equiv r$, $u_2\equiv\theta$, $u_3\equiv z$ son ampliamente usadas en aplicaciones y se obtienen de las cartesianas $x,y,z$ mediante la transformación (véase la figura 4.16)

$$\begin{aligned}r &= \sqrt{x^2+y^2}\\ \theta &= \arctan\frac{y}{x}\\ z &= z\end{aligned} \quad (4.112)$$

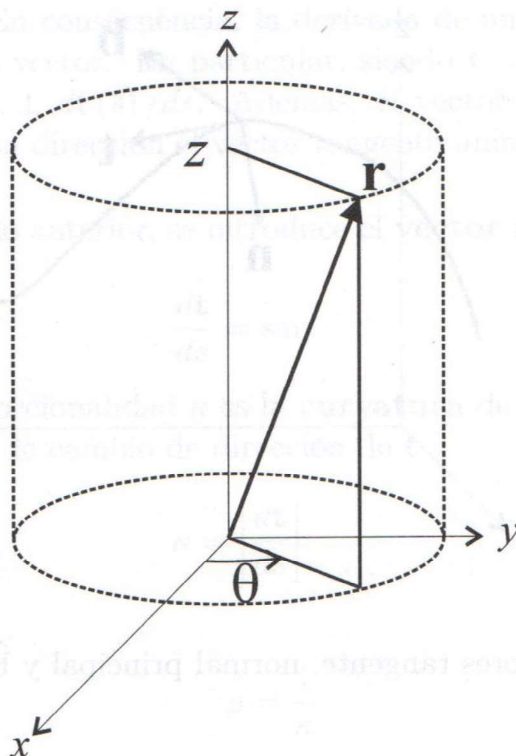



Figura 4.16. Coordenadas cilíndricas

El sistema de coordenadas cilíndricas es ortogonal y la longitud de arco infinitesimal es

$$ds^2 = dr^2 + r^2 d\theta^2 + dz^2 \tag{4.113}$$

De donde

$$h_1 = 1, \quad h_2 = r, \quad h_3 = 1 \tag{4.114}$$

En coordenadas cilíndricas, de (4.104), (4.109) y (4.111) se obtienen, respectivamente, el gradiente

$$\nabla f = \frac{\partial f}{\partial r}\mathbf{u}_r + \frac{1}{r}\frac{\partial f}{\partial \theta}\mathbf{u}_\theta + \frac{\partial f}{\partial z}\mathbf{u}_z \tag{4.115}$$

la divergencia

$$\nabla \cdot \mathbf{f} = \frac{1}{r}\left[\frac{\partial (r f_r)}{\partial r} + \frac{\partial f_\theta}{\partial \theta} + r\frac{\partial f_z}{\partial z}\right] \tag{4.116}$$

y laplaciano

$$\nabla^2 V = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial V}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 V}{\partial \theta^2} + \frac{\partial^2 V}{\partial z^2} \tag{4.117}$$

## 4.7. Tangente, curvatura y torsión. Fórmulas de Frenet-Serret

Suponemos que una curva $C$ en un espacio euclidiano de 3-D se representa mediante una función vectorial continuamente diferenciable,

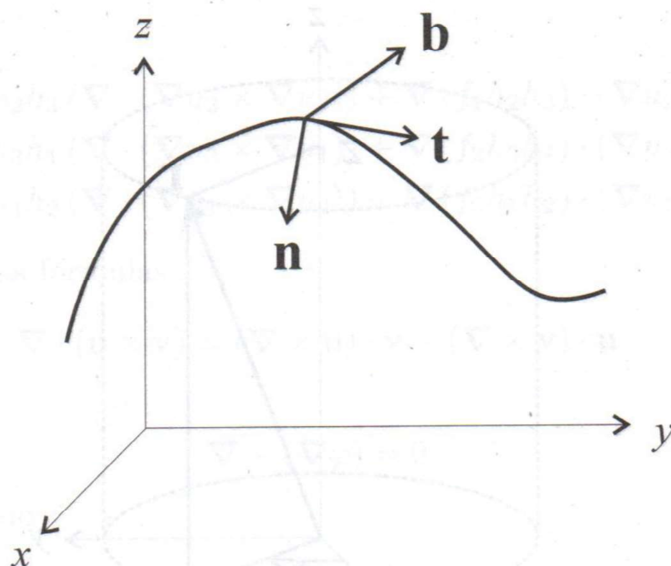



Figura 4.17. Vectores tangente, normal principal y binormal a una curva

$$\mathbf{r}(t) = x(t)\,\mathbf{i} + y(t)\,\mathbf{j} + z(t)\,\mathbf{k}$$

donde $t$ es un parámetro en un intervalo $t_1 \leq t \leq t_2$. El vector $d\mathbf{r}(t)/dt$ es un **vector tangente** a $C$ en el punto $P(t)$ y $\mathbf{u}(t) = \dot{\mathbf{r}}(t)/|\mathbf{r}(t)|$ es el **vector tangente unitario** (véase la figura 4.17). En particular, si $C$ se representa por $\mathbf{r}(s)$, donde $s$ es la longitud de arco, de la ecuación (4.77) se obtiene que

$$\frac{d\mathbf{r}(s)}{ds} \cdot \frac{d\mathbf{r}(s)}{ds} = \left(\frac{dx}{ds}\right)^2 + \left(\frac{dy}{ds}\right)^2 + \left(\frac{dz}{ds}\right)^2 = \frac{dx^2 + dy^2 + dz^2}{ds^2} = 1$$

Entonces, el vector

$$\mathbf{t} \equiv \mathbf{u}(s) = \frac{d\mathbf{r}(s)}{ds} \tag{4.118}$$

es el vector tangente unitario, $|\mathbf{t}| = 1$.

Es evidente que el vector de posición $\mathbf{q}$ de un punto $T$ en la recta tangente a la $C$ en el punto $P$ es la suma del vector de posición $\mathbf{r}$ de $P$ y un vector en la dirección de la tangente. De donde, una representación paramétrica de la tangente es

$$\mathbf{q}(\mu) = \mathbf{r}(t) + \mu \frac{d\mathbf{r}(t)}{dt} \tag{4.119}$$

Para continuar, demostramos una fórmula útil. Sea $\mathbf{a}(t)$ una función vectorial del parámetro $t$ que representa una curva $C$. Entonces,

$$\mathbf{a} \cdot \mathbf{a} = a^2 \Rightarrow \mathbf{a} \cdot \frac{d\mathbf{a}}{dt} = a\frac{da}{dt} \tag{4.120}$$

De la relación anterior se obtiene que

$$\mathbf{a} \cdot \frac{d\mathbf{a}}{dt} = 0 \quad \text{cuando} \quad |\mathbf{a}| = a = const \tag{4.121}$$

es decir, $(d\mathbf{a}/dt) \perp \mathbf{a}$. En consecuencia, la derivada de un vector "unitario" es perpendicular al mismo vector. En particular, siendo $\mathbf{t} = d\mathbf{r}(s)/ds$ un vector unitario, se tiene que $\mathbf{t} \perp d\mathbf{t}(s)/ds$. Además, el vector $d\mathbf{t}(s)/ds$ nos indica qué tan rápido cambia su dirección el vector tangente unitario $\mathbf{t}$ al moverse a lo largo de la curva $C$.

En consecuencia de lo anterior, se introduce el **vector normal principal n** mediante la relación

$$\frac{d\mathbf{t}}{ds} = \kappa \mathbf{n} \tag{4.122}$$

donde el factor de proporcionalidad $\kappa$ es la **curvatura** de $C$ en el punto $P$. La curvatura $\kappa$ es la razón de cambio de dirección de $\mathbf{t}$

$$\kappa = \left| \frac{d\mathbf{t}}{ds} \right| \tag{4.123}$$

La magnitud inversa a $\kappa$

$$\rho = \frac{1}{\kappa} \tag{4.124}$$

se llama *radio de curvatura* de $C$ en el punto $P$.

Definimos el **vector binormal** como

$$\mathbf{b} = \mathbf{t} \times \mathbf{n} \tag{4.125}$$

formando los vectores $\mathbf{t}$, $\mathbf{n}$ y $\mathbf{b}$ una base ortonormal local. Cualquier vector asociado con el punto $P$ de la curva $C$ puede representarse como combinación lineal de los tres vectores fundamentales $\mathbf{t}$, $\mathbf{n}$ y $\mathbf{b}$. Las tres rectas que pasan por el punto $P$ en las direcciones de $\mathbf{t}$, $\mathbf{n}$ y $\mathbf{b}$ son la **tangente**, la **normal principal** y la **binormal** de $C$.

Evaluaremos las derivadas $d\mathbf{b}(s)/ds$ y $d\mathbf{n}(s)/ds$. Dado que $\mathbf{b}$ es un vector unitario, su derivada es perpendicular a éste y, en consecuencia, se encuentra en el plano de vectores $\mathbf{t}$ y $\mathbf{n}$. Luego, como $\mathbf{b} \cdot \mathbf{t} = 0$ y $\mathbf{b} \cdot \mathbf{n} = 0$, se tiene

$$\frac{d}{ds}(\mathbf{b} \cdot \mathbf{t}) = \frac{d\mathbf{b}}{ds} \cdot \mathbf{t} + \mathbf{b} \cdot \frac{d\mathbf{t}}{ds} = \frac{d\mathbf{b}}{ds} \cdot \mathbf{t} = 0$$

Así que $(d\mathbf{b}/ds) \perp \mathbf{t}$ y, por lo tanto, $(d\mathbf{b}/ds) \parallel \mathbf{n}$. Entonces,

$$\frac{d\mathbf{b}}{ds} = \tau \mathbf{n} \tag{4.126}$$

donde el factor de proporcionalidad $\tau$ es la magnitud de $d\mathbf{b}/ds$. La función escalar $\tau(s)$ se llama *torsión* de la curva.

Para calcular $d\mathbf{n}/ds$, notamos que $\mathbf{b} = \mathbf{t} \times \mathbf{n}$, de donde $\mathbf{n} = \mathbf{b} \times \mathbf{t}$ (se le recuerda que $\mathbf{a} \times \mathbf{b} \times \mathbf{c} = (\mathbf{a} \cdot \mathbf{c})\,\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\,\mathbf{c}$). Luego,

$$\frac{d\mathbf{n}}{ds} = \frac{d}{ds}(\mathbf{b} \times \mathbf{t}) = \frac{d\mathbf{b}}{ds} \times \mathbf{t} + \mathbf{b} \times \frac{d\mathbf{t}}{ds} = \tau \mathbf{n} \times \mathbf{t} + \mathbf{b} \times \kappa \mathbf{n} = -\tau \mathbf{b} - \kappa \mathbf{t}$$

completándose con esta última el conjunto de las famosas fórmulas de Frenet-Serret:

$$\frac{d\mathbf{t}}{ds} = \kappa\mathbf{n} \tag{4.127}$$

$$\frac{d\mathbf{b}}{ds} = \tau\mathbf{n} \tag{4.128}$$

$$\frac{d\mathbf{n}}{ds} = -\left(\tau\mathbf{b}+\kappa\mathbf{t}\right) \tag{4.129}$$

**Ejemplo 4.3.** *Una hélice circular se describe por* $\mathbf{r}(t) = a\cos t\mathbf{i} + a\,\text{sen}\,t\mathbf{j} + bt\mathbf{k}$. *Se tiene que*

$$\mathbf{t} = \frac{d\mathbf{r}}{ds} = \frac{d\mathbf{r}}{dt}\frac{dt}{ds} = \left(-a\,\text{sen}\,t\mathbf{i} + a\cos t\mathbf{j} + b\mathbf{k}\right)\frac{dt}{ds}$$

$$\mathbf{t}\cdot\mathbf{t} = 1 = \left(a^2+b^2\right)\left(\frac{dt}{ds}\right)^2, \text{ de donde } \left(\frac{dt}{ds}\right)^2 = \left(a^2+b^2\right)^{-1}$$

*Así que*

$$\mathbf{t} = \left(-a\,\text{sen}\,t\mathbf{i} + a\cos t\mathbf{j} + b\mathbf{k}\right)\left(a^2+b^2\right)^{-1/2}$$

*Ahora,*

$$\kappa\mathbf{n} = \frac{d\mathbf{t}}{ds} = \frac{d\mathbf{t}}{dt}\frac{dt}{ds} = -a\left(\cos t\mathbf{i} + \text{sen}\,t\mathbf{j}\right)\left(a^2+b^2\right)^{-1}$$

*así que*

$$\mathbf{n} = -\left(\cos t\mathbf{i} + \text{sen}\,t\mathbf{j}\right), \quad \kappa = a\left(a^2+b^2\right)^{-1}$$

*Además,*

$$\mathbf{b} = \mathbf{t}\times\mathbf{n} = \left(b\,\text{sen}\,t\mathbf{i} - b\cos t\mathbf{j} + a\mathbf{k}\right)\left(a^2+b^2\right)^{-1/2}$$

$$\frac{d\mathbf{b}}{ds} = \tau\mathbf{n} = \left(b\cot s\mathbf{i} + b\,\text{sen}\,t\mathbf{j}\right)\left(a^2+b^2\right)^{-1}$$

*entonces,*

$$\tau = b\left(a^2+b^2\right)^{-1}$$

## 4.8. Planos fundamentales y ecuaciones intrínsecas

El plano que pasa por un punto $P$ de una curva $C$ y contiene a los vectores $\mathbf{t}$ y $\mathbf{n}$ se llama **plano osculante** a $C$ en el punto $P$; es perpendicular a $\mathbf{b}$. Sea $\mathbf{r}$ el vector de posición del punto $P$ en la curva $C$ y sea $\mathbf{s}$ un vector variable de posición de los puntos en el plano osculante. Entonces, la ecuación del plano osculante se puede escribir en la forma

$$(\mathbf{s}-\mathbf{r})\cdot\mathbf{b} = 0 \tag{4.130}$$

El **plano normal** a la curva $C$ que pasa por un punto $P$ se define como el plano que contiene a $\mathbf{n}$ y $\mathbf{b}$ y, por tanto, es perpendicular a $\mathbf{t}$. La ecuación del plano normal es

$$(\mathbf{s}-\mathbf{r})\cdot\mathbf{t}=0 \tag{4.131}$$

El **plano rectificador** a la curva $C$ que pasa por un punto $P$ se define como el plano que contiene a $\mathbf{t}$ y $\mathbf{b}$ y, por tanto, es perpendicular a $\mathbf{n}$. La ecuación del plano rectificador es

$$(\mathbf{s}-\mathbf{r})\cdot\mathbf{n}=0 \tag{4.132}$$

La curvatura y torsión de una curva dependen de la posición del punto $P$ en la curva y, en consecuencia, dependen del parámetro $s$. Sea

$$\kappa=\kappa(s)\,,\quad \tau=\tau(s) \tag{4.133}$$

Estas dos ecuaciones reciben el nombre de **ecuaciones intrínsecas** de la curva. El nombre se debe al hecho de que dos curvas con las mismas ecuaciones intrínsecas son idénticas, excepto, posiblemente, por la orientación en el espacio. Suponemos que dos curvas, $C_1$ y $C_2$, tienen las mismas ecuaciones intrínsecas

$$\kappa(s)\equiv\kappa_1(s)=\kappa_2(s)\,,\quad \tau(s)\equiv\tau_1(s)=\tau_2(s)$$

Si es necesario, hacemos coincidir las ternas $\mathbf{t}$, $\mathbf{n}$ y $\mathbf{b}$ en un punto correspondiente $P(s)$ mediante un movimiento rígido de una de las curvas. Para cualquier valor del parámetro $s$, se tiene

$$\begin{aligned}
\frac{d}{ds}(\mathbf{t}_1\cdot\mathbf{t}_2) &= \mathbf{t}_1\cdot\kappa\mathbf{n}_2+\kappa\mathbf{n}_1\cdot\mathbf{t}_2\\
\frac{d}{ds}(\mathbf{n}_1\cdot\mathbf{n}_2) &= \mathbf{n}_1\cdot(-\kappa\mathbf{t}_2-\tau\mathbf{b}_2)+(-\kappa\mathbf{t}_1-\tau\mathbf{b}_1)\cdot\mathbf{n}_2\\
\frac{d}{ds}(\mathbf{b}_1\cdot\mathbf{b}_2) &= \mathbf{b}_1\cdot\tau\mathbf{n}_2+\tau\mathbf{n}_1\cdot\mathbf{b}_2
\end{aligned}$$

Sumando las ecuaciones anteriores, se obtiene que

$$\frac{d}{ds}(\mathbf{t}_1\cdot\mathbf{t}_2+\mathbf{n}_1\cdot\mathbf{n}_2+\mathbf{b}_1\cdot\mathbf{b}_2)=0$$

para todo valor del parámetro $s$. Dado que en el punto $P$ es válida la igualdad $\mathbf{t}_1\cdot\mathbf{t}_2+\mathbf{n}_1\cdot\mathbf{n}_2+\mathbf{b}_1\cdot\mathbf{b}_2=3$, de la ecuación anterior se concluye que la igualdad

$$\mathbf{t}_1\cdot\mathbf{t}_2+\mathbf{n}_1\cdot\mathbf{n}_2+\mathbf{b}_1\cdot\mathbf{b}_2=3$$

es válida para todo valor del parámetro $s$, esto es,

$$\mathbf{t}_1=\mathbf{t}_2,\quad \mathbf{n}_1=\mathbf{n}_2,\quad \mathbf{b}_1=\mathbf{b}_2$$

$$\frac{d}{ds}(\mathbf{r}_1)=\frac{d}{ds}(\mathbf{r}_2)\quad \text{o}\quad \mathbf{r}_1(s)=\mathbf{r}_2(s)$$

Por lo tanto, las dos curvas son idénticas en una vecindad del punto $P$ y, por suposición de la analiticidad de las curvas, éstas son idénticas en su totalidad.

## 4.9. Involuta, evoluta y envolvente

### 4.9.1. Involuta

Considérese una curva espacial $\Gamma$. La **involuta** $\Gamma_1$ de la curva $\Gamma$ se define como cualquier curva que es normal a todas las tangentes de $\Gamma$ (véase la figura 4.18). De la figura 4.18 es evidente que la ecuación de la involuta es

$$\mathbf{r}_1 = \mathbf{r} + u\mathbf{t} \tag{4.134}$$

en donde $u$ es un factor desconocido de proporcionalidad entre los vectores $\mathbf{r}_1 - \mathbf{r}$ y $\mathbf{t}$, $\mathbf{r}$ es el vector de posición de un punto $P$ sobre la curva $\Gamma$ y $\mathbf{r}_1$ es el vector de posición del punto de intersección de $\Gamma_1$ con la tangente en el punto $P$ a la curva $\Gamma$. Al diferenciar (4.134) se obtiene

$$\frac{d\mathbf{r}_1}{ds_1} = \mathbf{t}_1 = \left(\frac{d\mathbf{r}}{ds} + \frac{du}{ds}\mathbf{t} + u\kappa\mathbf{n}\right)\frac{ds}{ds_1} \tag{4.135}$$

en donde $s$ y $s_1$ son la longitud de arco a lo largo de $\Gamma$ y $\Gamma_1$, respectivamente. De (4.135) y la definición de la involuta se tiene

$$0 = \mathbf{t}_1 \cdot \mathbf{t} = \left[\left(\mathbf{t} + \frac{du}{ds}\mathbf{t} + u\kappa\mathbf{n}\right)\frac{ds}{ds_1}\right] \cdot \mathbf{t} = 1 + \frac{du}{ds} \tag{4.136}$$

De esta ecuación se obtiene que

$$u = c - s \tag{4.137}$$

en donde $c$ es una constante. Entonces, la ecuación (4.134) de $\Gamma_1$ se escribe como

$$\mathbf{r}_1 = \mathbf{r} + (c - s)\,\mathbf{t} \tag{4.138}$$

y, en consecuencia, para $\Gamma$ existe una familia infinita de involutas; cada involuta corresponde a un valor de constante $c$. La distancia entre cada par de involutas se mantiene constante. Una involuta puede trazarse mediante el extremo de una cuerda no extensible de longitud $c$ que se enrolla a lo largo de $\Gamma$ de tal manera que en cada instante apunte a la dirección tangente de $\Gamma$ (véase la figura 4.18). De esta manera queda claro que $\Gamma \perp \Gamma_1$ en el punto de intersección de éstas, cuando $s = c$.

De (4.122) y (4.138), se tiene

$$\begin{aligned}\mathbf{t}_1 &= \frac{d\mathbf{r}_1}{ds_1} = \left(\frac{d\mathbf{r}}{ds} + (c - s)\frac{d\mathbf{t}}{ds} - \mathbf{t}\right)\frac{ds}{ds_1} \\ &= (c - s)\,\kappa\mathbf{n}\frac{ds}{ds_1}\end{aligned} \tag{4.139}$$

es decir, la tangente a la involuta $\Gamma_1$ es paralela a la normal correspondiente de $\Gamma$. Dado que $\mathbf{t}_1$ y $\mathbf{n}$ son vectores unitarios, de (4.139) se obtiene la igualdad

$$(c - s)\,\kappa\frac{ds}{ds_1} = 1.$$

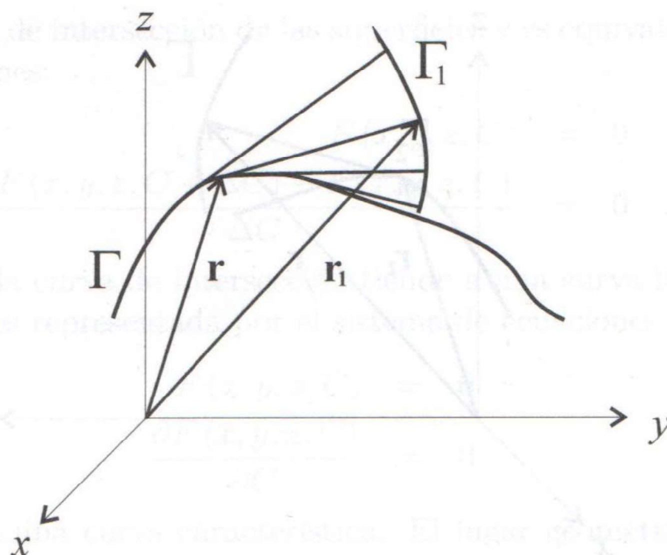



Figura 4.18. Involuta

Por consiguiente, la curvatura de la involuta $\kappa_1$ se obtiene mediante la relación

$$\frac{d\mathbf{t}_1}{ds_1} = \kappa_1\mathbf{n}_1 = \frac{d\mathbf{n}}{ds_1} = \frac{d\mathbf{n}}{ds}\frac{ds}{ds_1} = \frac{(-\kappa\mathbf{t} - \tau\mathbf{b})}{(c-s)\,\kappa}$$

De donde

$$\kappa_1^2 = \frac{\kappa^2 + \tau^2}{(c-s)^2\,\kappa^2} \tag{4.140}$$

### 4.9.2. Evoluta

Una curva $\Gamma_2$ cuyas tangentes son perpendiculares a la curva dada $\Gamma$, es llamada la **evoluta** de la curva $\Gamma$. La tangente a $\Gamma_2$ debe estar en el plano de los vectores **n** y **b**, dado que es perpendicular a **t** (véase la figura 4.19). En consecuencia, la ecuación de la evoluta es

$$\mathbf{r}_2 = \mathbf{r} + u\mathbf{n} + v\mathbf{b} \tag{4.141}$$

Al diferenciarla, se tiene

$$\begin{aligned}\mathbf{t}_2 &= \frac{d\mathbf{r}_2}{ds_2} = \left[\frac{d\mathbf{r}}{ds} + u\frac{d\mathbf{n}}{ds} + v\frac{d\mathbf{b}}{ds} + \frac{du}{ds}\mathbf{n} + \frac{dv}{ds}\mathbf{b}\right]\frac{ds}{ds_2} \\ &= \left[\mathbf{t} + u\,(-\kappa\mathbf{t} - \tau\mathbf{b}) + v\tau\mathbf{n} + \frac{du}{ds}\mathbf{n} + \frac{dv}{ds}\mathbf{b}\right]\frac{ds}{ds_2}\end{aligned}$$

Es claro que $\mathbf{t}_2 \cdot \mathbf{t} = 0$, de donde $0 = \mathbf{t}_2 \cdot \mathbf{t} = 1 - u\kappa$ y $u = 1/\kappa = \rho$ es el radio de curvatura de $\Gamma$. Entonces,

$$\mathbf{t}_2 = \left[\left(\frac{du}{ds} + v\tau\right)\mathbf{n} + \left(\frac{dv}{ds} - u\tau\right)\mathbf{b}\right]\frac{ds}{ds_2}$$

Por otro lado, $\mathbf{t}_2$ es paralelo a $\mathbf{r}_2 - \mathbf{r} = u\mathbf{n} + v\mathbf{b}$, lo que implica

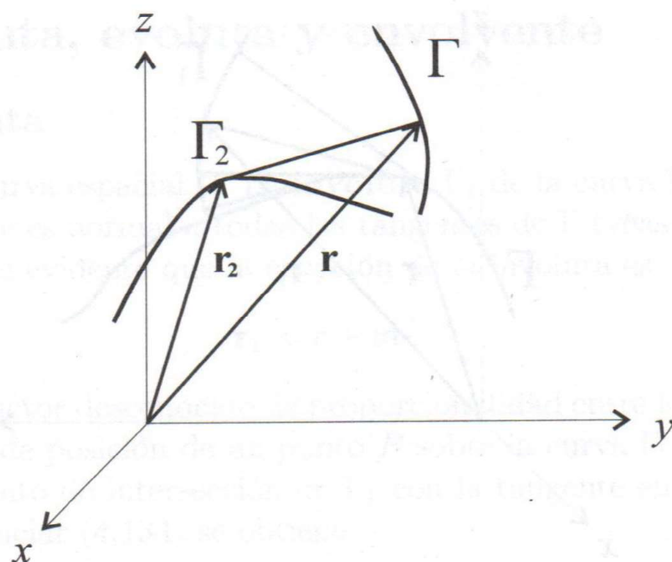



Figura 4.19. Evoluta

$$\left(\frac{du}{ds}+v\tau\right)/u=\left(\frac{dv}{ds}-u\tau\right)/v$$

De donde

$$\tau=\frac{v'u-u'v}{u^2+v^2}=\frac{d}{ds}\left(\arctan\frac{v}{u}\right)$$

$$\varphi\equiv\int_0^s\tau ds=\arctan\frac{v}{u}+C \tag{4.142}$$

y $v=\rho\tan(\varphi-C)$, ya que $u=\rho$. Finalmente, de (4.141) se obtiene la ecuación

$$\mathbf{r}_2=\mathbf{r}+\rho\mathbf{n}+\rho\tan(\varphi-C)\,\mathbf{b} \tag{4.143}$$

que, en realidad, es una familia uniparamétrica de evolutas de la curva Γ.

## 4.9.3. Envolvente

La ecuación

$$F(x,y,z,C)=0 \tag{4.144}$$

en donde $C$ es un parámetro, representa una familia uniparamétrica de superficies. Una superficie vecina es

$$F(x,y,z,C+\Delta C)=0$$

El sistema de ecuaciones

$$\begin{aligned}F(x,y,z,C)&=0\\F(x,y,z,C+\Delta C)&=0\end{aligned}$$

representa la curva de intersección de las superficies y es equivalente al siguiente sistema de ecuaciones:

$$\begin{aligned} F(x,y,z,C) &= 0 \\ \frac{F(x,y,z,C+\Delta C)-F(x,y,z,C)}{\Delta C} &= 0 \end{aligned}$$

Cuando $\Delta C \to 0$, la curva de intersección tiende a una curva llamada la curva característica que es representada por el sistema de ecuaciones

$$\begin{aligned} F(x,y,z,C) &= 0 \\ \frac{\partial F(x,y,z,C)}{\partial C} &= 0 \end{aligned}$$

Cada $C$ determina una curva característica. El lugar geométrico de todas las curvas características, que se obtiene eliminando a $C$ en este último sistema, es una superficie llamada la **envolvente** de la familia uniparamétrica de superficies (4.144).

# 4.10. Superficies y coordenadas curvilíneas

Considérese el sistema de ecuaciones

$$x = x(u,v), \quad y = y(u,v), \quad z = z(u,v) \tag{4.145}$$

en donde $x, y, z$ son coordenadas en un sistema cartesiano, $u$ y $v$ son variables reales. En general, el sistema de ecuaciones (4.145) representa una superficie $S$ en $3D$. Si $v$ se mantiene fija, las ecuaciones (4.145) representan una curva espacial que es llamada la curva-$u$. Si $u$ se mantiene fija, variando $v$ tenemos una curva espacial que es llamada la curva-$v$. Dichas curvas se encuentran sobre la superficie $S$ y, por tanto, $u$ y $v$ pueden considerarse como coordenadas curvilíneas sobre $S$. Las ecuaciones (4.145) pueden escribirse como

$$\mathbf{r}(u,v) = x(u,v)\,\mathbf{i} + y(u,v)\,\mathbf{j} + z(u,v)\,\mathbf{k} \tag{4.146}$$

donde $\mathbf{r}(u,v)$ es el vector indicador de los puntos de $S$.

## 4.10.1. Longitud de arco sobre una superficie y primera forma fundamental

Sean $\mathbf{r}(u,v)$ y $\mathbf{r}(u,v) + d\mathbf{r}(u,v)$ dos puntos cercanos en $S$. La distancia entre estos puntos es

$$\begin{aligned} ds^2 &= d\mathbf{r}\cdot d\mathbf{r} = \left(\frac{\partial \mathbf{r}}{\partial u}du + \frac{\partial \mathbf{r}}{\partial v}dv\right)^2 \\ &= \left(\frac{\partial \mathbf{r}}{\partial u}\right)^2 du^2 + 2\left(\frac{\partial \mathbf{r}}{\partial u}\frac{\partial \mathbf{r}}{\partial v}\right) dudv + \left(\frac{\partial \mathbf{r}}{\partial v}\right)^2 dv^2 \\ &\equiv Edu^2 + 2Fdudv + Gdv^2 \end{aligned} \tag{4.147}$$

Esta última ecuación es conocida como **primera forma fundamental**, en donde

$$E(u,v)=\left(\frac{\partial \mathbf{r}}{\partial u}\right)^2, \quad F(u,v)=\frac{\partial \mathbf{r}}{\partial u}\frac{\partial \mathbf{r}}{\partial v}, \quad G(u,v)=\left(\frac{\partial \mathbf{r}}{\partial v}\right)^2 \tag{4.148}$$

En particular, sobre una curva-$u$, $dv=0$ y

$$ds_u=\sqrt{E}du \tag{4.149}$$

mientras que sobre una curva-$v$, $du=0$

$$ds_v=\sqrt{G}dv \tag{4.150}$$

Los vectores

$$\frac{\partial \mathbf{r}}{\partial u}=\frac{\partial \mathbf{r}}{\partial s_u}\frac{ds_u}{du}=\sqrt{E}\frac{\partial \mathbf{r}}{\partial s_u} \quad \text{y} \quad \frac{\partial \mathbf{r}}{\partial v}=\frac{\partial \mathbf{r}}{\partial s_v}\frac{ds_v}{du}=\sqrt{G}\frac{\partial \mathbf{r}}{\partial s_v} \tag{4.151}$$

son tangentes a las curvas -$u$ y -$v$, respectivamente. Además,

$$\mathbf{t}_u=\frac{\partial \mathbf{r}}{\partial s_u} \quad \text{y} \quad \mathbf{t}_v=\frac{\partial \mathbf{r}}{\partial s_v} \tag{4.152}$$

son vectores tangentes unitarios, $|\mathbf{t}_u|=|\mathbf{t}_v|=1$. Entonces, las curvas paramétricas -$u$ y -$v$ forman un sistema ortogonal de coordenadas en $S$ si y sólo si

$$\frac{\partial \mathbf{r}}{\partial u}\cdot\frac{\partial \mathbf{r}}{\partial v}=F(u,v)=0 \tag{4.153}$$

**Ejemplo 4.4.** *Considérese una superficie*

$$\mathbf{r}(\theta,\varphi)=r\operatorname{sen}\theta\cos\varphi\mathbf{i}+r\operatorname{sen}\theta\operatorname{sen}\varphi\mathbf{j}+r\cos\theta\mathbf{k}$$

*donde $r$ es una constante. Al diferenciar, se tiene*

$$\frac{\partial \mathbf{r}}{\partial \theta}=r\cos\theta\cos\varphi\hat{\mathbf{i}}+r\cos\theta\operatorname{sen}\varphi\hat{\mathbf{j}}-r\operatorname{sen}\theta\hat{\mathbf{k}}$$

$$\frac{\partial \mathbf{r}}{\partial \varphi}=-r\operatorname{sen}\theta\operatorname{sen}\varphi\hat{\mathbf{i}}+r\operatorname{sen}\theta\cos\varphi\hat{\mathbf{j}}$$

*Luego,*

$$E=\left(\frac{\partial \mathbf{r}}{\partial \theta}\right)^2=r^2$$

$$F=\frac{\partial \mathbf{r}}{\partial \theta}\frac{\partial \mathbf{r}}{\partial \varphi}=0$$

$$G=\left(\frac{\partial \mathbf{r}}{\partial \varphi}\right)^2=r^2\operatorname{sen}^2\theta$$

*Entonces,*

$$ds^2=r^2d\theta^2+r^2\operatorname{sen}^2\theta d\varphi^2$$

*y las curvas -$\theta$ y -$\varphi$ son ortogonales entre sí. Por supuesto, la superficie es una esfera.*

Como se verá en la siguiente sección, la primera forma fundamental tiene una importancia básica porque permite medir longitudes de curvas, ángulos entre las curvas y áreas sobre la superficie $S$.

## 4.10.2. Curvas en superficie

Sea $\mathbf{r} = \mathbf{r}(u, v)$ la ecuación de una superficie $S$. Entonces, una curva $\Gamma$ en la superficie $S$ se define mediante la ecuación

$$\mathbf{r} = \mathbf{r}(u(t), v(t)) \tag{4.154}$$

en donde los parámetros $u$ y $v$ son funciones de una sola variable $t$. Cada par de funciones $u(t)$ y $v(t)$ determinan una curva en $S$. Por ejemplo, para una curva-$u$, $u \equiv t$ y $v$ es una constante.

Sean $\Gamma$ y $\Gamma'$ dos curvas en $S$ definidas por $\mathbf{r} = \mathbf{r}(u(t), v(t))$ y $\mathbf{r} = \mathbf{r}(u'(t), v'(t))$, respectivamente. A lo largo de la curva $\Gamma$,

$$d\mathbf{r}_\Gamma = \left(\frac{\partial \mathbf{r}}{\partial u}\frac{\partial u}{\partial t} + \frac{\partial \mathbf{r}}{\partial v}\frac{\partial v}{\partial t}\right) dt = \frac{\partial \mathbf{r}}{\partial u} du + \frac{\partial \mathbf{r}}{\partial v} dv$$

mientras que a lo largo de $\Gamma'$,

$$d\mathbf{r}_{\Gamma'} = \left(\frac{\partial \mathbf{r}}{\partial u}\frac{\partial u'}{\partial t} + \frac{\partial \mathbf{r}}{\partial v}\frac{\partial v'}{\partial t}\right) dt = \frac{\partial \mathbf{r}}{\partial u} du' + \frac{\partial \mathbf{r}}{\partial v} dv'$$

Los vectores $d\mathbf{r}_\Gamma$ y $d\mathbf{r}_{\Gamma'}$ son completamente determinados por las pares $(du, dv)$ y $(du', dv')$. El producto escalar de dos vectores diferenciales es

$$d\mathbf{r}_\Gamma \cdot d\mathbf{r}_{\Gamma'} = E(u, v)\, du du' + F(u, v)(dv du' + du dv') + G(u, v)\, dv dv'$$

Entonces, las dos curvas $\Gamma$ y $\Gamma'$ son perpendiculares entre sí en el punto de su intersección si y sólo si

$$d\mathbf{r}_\Gamma \cdot d\mathbf{r}_{\Gamma'} = E du du' + F(dv du' + du dv') + G dv dv' = 0$$

es decir,

$$E + F\left(\frac{dv}{du} + \frac{dv'}{du'}\right) + G\left(\frac{dv}{du}\frac{dv'}{du'}\right) = 0 \tag{4.155}$$

Si en $S$ se tiene un sistema de curvas definidas por la ecuación diferencial $P(u, v)\, du' + Q(u, v)\, dv' = 0$, entonces la ecuación diferencial para las trayectorias ortogonales se da por

$$E + F\left(\frac{dv}{du} - \frac{P}{Q}\right) + \frac{GP}{Q}\frac{dv}{du} = 0 \tag{4.156}$$

ya que $dv'/du' = -P(u, v)/Q(u, v)$.

**Ángulo**

El ángulo entre dos curvas, $\Gamma$ y $\Gamma'$, sobre la superficie $S$ en un punto de intersección de éstas se calcula de la siguiente manera. Sean $\Gamma$ y $\Gamma'$ dos curvas en $S$ definidas por $\mathbf{r} = \mathbf{r}(u(t), v(t))$ y $\mathbf{r} = \mathbf{r}(u'(t), v'(t))$, respectivamente, las cuales se intersectan en un punto $P$ de $S$. Los vectores

$$\mathbf{t} = \frac{d}{dt}\mathbf{r}(u(t), v(t)) = \frac{\partial \mathbf{r}}{\partial u}\frac{\partial u}{\partial t} + \frac{\partial \mathbf{r}}{\partial v}\frac{\partial v}{\partial t} \tag{4.157}$$

y

$$\mathbf{t}' = \frac{d}{dt}\mathbf{r}(u'(t), v'(t)) = \frac{\partial \mathbf{r}}{\partial u}\frac{\partial u'}{\partial t} + \frac{\partial \mathbf{r}}{\partial v}\frac{\partial v'}{\partial t} \tag{4.158}$$

son tangentes a $\Gamma$ y $\Gamma'$, respectivamente. El ángulo entre $\Gamma$ y $\Gamma'$ en $P$ se define como el ángulo entre $\mathbf{t}$ y $\mathbf{t}'$ y, por tanto,

$$\cos\theta = \frac{\mathbf{t}\cdot\mathbf{t}'}{|\mathbf{t}|\,|\mathbf{t}'|} \tag{4.159}$$

en donde

$$\begin{aligned}\mathbf{t}\cdot\mathbf{t}' &= \left(\frac{\partial \mathbf{r}}{\partial u}\frac{\partial u}{\partial t} + \frac{\partial \mathbf{r}}{\partial v}\frac{\partial v}{\partial t}\right)\cdot\left(\frac{\partial \mathbf{r}}{\partial u}\frac{\partial u'}{\partial t} + \frac{\partial \mathbf{r}}{\partial v}\frac{\partial v'}{\partial t}\right)\\ &= E\frac{\partial u}{\partial t}\frac{\partial u'}{\partial t} + F\left(\frac{\partial u}{\partial t}\frac{\partial v'}{\partial t} + \frac{\partial u'}{\partial t}\frac{\partial v}{\partial t}\right) + G\frac{\partial v}{\partial t}\frac{\partial v'}{\partial t}\end{aligned}$$

y

$$\begin{aligned}|\mathbf{t}| &= \left(E\left(\frac{\partial u}{\partial t}\right)^2 + 2F\frac{\partial u}{\partial t}\frac{\partial v}{\partial t} + G\left(\frac{\partial v}{\partial t}\right)^2\right)^{1/2}\\ |\mathbf{t}'| &= \left(E\left(\frac{\partial u'}{\partial t}\right)^2 + 2F\frac{\partial u'}{\partial t}\frac{\partial v'}{\partial t} + G\left(\frac{\partial v'}{\partial t}\right)^2\right)^{1/2}\end{aligned}$$

Este resultado muestra que el ángulo entre dos curvas de una superficie en un punto de su intersección se puede expresar en función de $E$, $F$, $G$, que son características propias de la superficie, y de las derivadas de las funciones que representan las curvas.

**Longitud**

Basándose en (4.73) y (4.157) se obtiene que una curva $\Gamma$, definida por $\mathbf{r} = \mathbf{r}(u(t), v(t))$ $(a < t < b)$, tiene la longitud

$$L = \int_a^b \sqrt{\dot{\mathbf{r}}\cdot\dot{\mathbf{r}}}\,dt = \int_a^b \left(E\left(\frac{\partial u}{\partial t}\right)^2 + 2F\frac{\partial u}{\partial t}\frac{\partial v}{\partial t} + G\left(\frac{\partial v}{\partial t}\right)^2\right)^{1/2} dt \tag{4.160}$$

**Área**

Para calcular el área de un dominio $D$ de la superficie $S$ se procede de la siguiente manera. Dado que $\partial\mathbf{r}/\partial u$ es tangente a la curva-$u$ y $\partial\mathbf{r}/\partial v$ es tangente a la curva-$v$, la diferencial de área es

$$dS\left(u,v\right)=\left|d\mathbf{r}_u\times d\mathbf{r}_v\right|=\left|\frac{\partial\mathbf{r}}{\partial u}\right|\left|\frac{\partial\mathbf{r}}{\partial v}\right|\operatorname{sen}\theta dudv=\sqrt{EG}\operatorname{sen}\theta dudv$$

donde $d\mathbf{r}_u=\left(\partial\mathbf{r}/\partial u\right)du$ y $d\mathbf{r}_v=\left(\partial\mathbf{r}/\partial v\right)dv$ y hemos usado (4.151) y (4.152). También

$$F=\left|\frac{\partial\mathbf{r}}{\partial u}\right|\left|\frac{\partial\mathbf{r}}{\partial v}\right|\cos\theta=\sqrt{EG}\cos\theta$$

De donde

$$dS\left(u,v\right)=\sqrt{EG}\operatorname{sen}\theta dudv=\sqrt{EG}\sqrt{1-\frac{F^2}{EG}}dudv=\sqrt{EG-F^2}dudv$$

Por lo tanto,

$$S_D=\int_D dS\left(u,v\right)=\int_D\sqrt{EG-F^2}dudv \tag{4.161}$$

**Problema.** *En la superficie* $\mathbf{r}=u\mathbf{i}+v\mathbf{j}$, *encuentre las trayectorias ortogonales a las curvas* $uv=const.$

## 4.10.3. Plano tangente y normal a una superficie

Sea $S$ una superficie suave representada por una función vectorial $\mathbf{r}\left(u,v\right)$ y $\Gamma$ una curva sobre $S$ representada por (4.154). Supóngase que las dos funciones $u\left(t\right)$ y $v\left(t\right)$ en (4.154) tienen primeras derivadas continuas para cada $t$ y por lo menos una de éstas no es cero. Entonces $\Gamma$ tiene una tangente en cada uno de sus puntos y un vector tangente es

$$\frac{d\mathbf{r}\left(u\left(t\right),v\left(t\right)\right)}{dt}=\frac{\partial\mathbf{r}}{\partial u}\frac{du}{dt}+\frac{\partial\mathbf{r}}{\partial v}\frac{dv}{dt} \tag{4.162}$$

Si los vectores $\partial\mathbf{r}/du$ y $\partial\mathbf{r}/dv$ son linealmente independientes, es decir, las curvas -$u$ y -$v$ no son colineales en ningún punto, entonces éstos determinan un plano. Este plano se llama **plano tangente** a $S$ en el punto correspondiente de ésta. De (4.162) se concluye que el plano tangente contiene a la tangente a cualquier curva sobre $S$ que pasa por el punto $P$, en ese punto.

La perpendicular al plano tangente en un punto $P$ de $S$ se llama **normal a la superficie** $S$ en el punto $P$. Puesto que los vectores $\partial\mathbf{r}/\partial u$ y $\partial\mathbf{r}/\partial v$ son tangentes a la superficie $S$ a lo largo de curva-$u$ y curva-$v$, respectivamente, la normal unitaria a la superficie $S$ es

$$\begin{aligned}\mathbf{n} &= \frac{\frac{\partial\mathbf{r}}{\partial u}\times\frac{\partial\mathbf{r}}{\partial v}}{\left|\frac{\partial\mathbf{r}}{\partial u}\times\frac{\partial\mathbf{r}}{\partial v}\right|}=\left(\sqrt{EG-F^2}\right)^{-1}\left(\frac{\partial\mathbf{r}}{\partial u}\times\frac{\partial\mathbf{r}}{\partial v}\right) \qquad (4.163)\\ &= \sqrt{EG}\left(\sqrt{EG-F^2}\right)^{-1}\left(\mathbf{t}_u\times\mathbf{t}_v\right)\end{aligned}$$

### 4.10.4. Segunda forma fundamental

Considérense todos los planos que contienen la normal $\mathbf{n}$ a una superficie $\mathbf{r} = \mathbf{r}(u, v)$ en un punto $P$. Las intersecciones de estos planos con dicha superficie es una familia de curvas tales que $\mathbf{n}$ también es una normal a cada una de estas curvas. Sea $\Gamma$ una curva de éstas sobre la superficie $S$. La ecuación vectorial de la curva sea $\mathbf{r} = \mathbf{r}(u(s), v(s))$, en donde $s$ es el parámetro natural, es decir, $ds$ es la longitud de arco a lo largo de $\Gamma$. El vector tangente unitario a $\Gamma$ es

$$\mathbf{t} = \frac{d\mathbf{r}}{ds} = \frac{\partial \mathbf{r}}{\partial u}\frac{du}{ds} + \frac{\partial \mathbf{r}}{\partial v}\frac{dv}{ds}$$

y, por lo tanto,

$$\begin{aligned}\frac{d\mathbf{t}}{ds} &= \kappa_n \mathbf{n} \\ &= \frac{\partial^2 \mathbf{r}}{\partial u^2}\left(\frac{du}{ds}\right)^2 + \left(\frac{\partial^2 \mathbf{r}}{\partial v \partial u} + \frac{\partial^2 \mathbf{r}}{\partial u \partial v}\right)\frac{\partial u}{\partial s}\frac{dv}{ds} + \\ &\quad \frac{\partial^2 \mathbf{r}}{\partial v^2}\left(\frac{dv}{ds}\right)^2 + \frac{\partial \mathbf{r}}{\partial u}\frac{d^2 u}{ds^2} + \frac{\partial \mathbf{r}}{\partial v}\frac{d^2 v}{ds^2}\end{aligned}$$

De donde la curvatura

$$\begin{aligned}\kappa_n &= \mathbf{n} \cdot (\kappa_n \mathbf{n}) \\ &= \left(\mathbf{n} \cdot \frac{\partial^2 \mathbf{r}}{\partial u^2}\right)\left(\frac{du}{ds}\right)^2 + 2\left(\mathbf{n}\frac{\partial^2 \mathbf{r}}{\partial u \partial v}\right)\frac{\partial u}{\partial s}\frac{dv}{ds} + \left(\mathbf{n} \cdot \frac{\partial^2 \mathbf{r}}{\partial v^2}\right)\left(\frac{dv}{ds}\right)^2\end{aligned}$$

ya que $\mathbf{n} \cdot \partial \mathbf{r}/du = \mathbf{n} \cdot \partial \mathbf{r}/dv = 0$. Entonces,

$$\begin{aligned}\kappa_n &= \frac{edu^2 + 2f\,dudv + gdv^2}{ds^2} \\ &= \frac{edu^2 + 2f\,dudv + gdv^2}{Edu^2 + 2F\,dudv + Gdv^2}\end{aligned} \qquad (4.164)$$

donde

$$e = \mathbf{n}\frac{\partial^2 \mathbf{r}}{\partial u^2}, \qquad f = \mathbf{n} \cdot \frac{\partial^2 \mathbf{r}}{\partial u \partial v}, \qquad g = \mathbf{n}.\frac{\partial^2 \mathbf{r}}{\partial v^2} \qquad (4.165)$$

La magnitud

$$edu^2 + 2f\,dudv + gdv^2 \qquad (4.166)$$

se llama la **segunda forma fundamental** de $S$.

Para comprender el significado geométrico de la segunda forma fundamental, trazamos el plano tangencial a la superficie $S$ en un punto $\mathbf{r}_0 = \mathbf{r}(u_0, v_0)$. La distancia $D$ entre un punto vecino $\mathbf{r}(u_0 + du, v_0 + dv)$ de $S$ y el plano tangencial en el punto $\mathbf{r}(u_0, v_0)$ es igual a $D = \Delta\mathbf{r} \cdot \mathbf{n}$, donde

$$\Delta\mathbf{r} = \mathbf{r} - \mathbf{r}_0 = \frac{\partial \mathbf{r}}{\partial u}du + \frac{\partial \mathbf{r}}{\partial v}dv + \frac{1}{2!}\left(\frac{\partial^2 \mathbf{r}}{\partial u^2} + 2\frac{\partial^2 \mathbf{r}}{\partial u \partial v} + \frac{\partial^2 \mathbf{r}}{\partial v^2}\right) + \ldots$$

es el desarrollo en serie de Taylor. Entonces,

$$D = \Delta \mathbf{r} \cdot \mathbf{n} = \frac{1}{2!}\left(edu^2 + 2fdudv + gdv^2\right) \qquad (4.167)$$

salvo los términos infinitesimales de orden mayor.

**Nota bibliográfica:** para ampliar los conocimientos sobre el tema de este capítulo, puede consultar [1], [11], [13], [14] y [17].

## 10.4 Segunda forma fundamental

[illegible]

$$eda^2 + 2f\,du\,dv + g\,dv^2$$

[illegible] la segunda forma fundamental de $S$.

[illegible] de la segunda forma [illegible] en un punto [illegible] $+ dv)$ de $S$ y [illegible] donde

[illegible]

# Capítulo 5

# Espacios vectoriales

El conjunto de los vectores, junto con ciertas operaciones algebraicas sobre éstos, sugiere una estructura algebraica que se conoce como espacio vectorial (o espacio lineal). El concepto de espacio lineal es importante porque permite tratar, de manera uniforme y generalizada, muchos conjuntos (conjuntos de vectores, matrices, funciones, operadores, etc.) que se presentan con frecuencia en las Matemáticas y sus aplicaciones. De hecho, este concepto es básico en el análisis funcional (análisis moderno abstracto), en las ecuaciones diferenciales, en el análisis numérico y en el análisis de sistemas y otros campos de interés práctico para la Física e Ingeniería. Cabe mencionar que el formalismo matemático de la Mecánica Cuántica está esencialmente basado en el concepto de espacios lineales.

Suponiendo que el estudiante está familiarizado con lo básico sobre los vectores, en este capítulo se da una introducción al formalismo de espacios vectoriales.

## 5.1. Espacio lineal, base y coordenadas

### 5.1.1. Postulados

Cualquier conjunto no vacío $S$ de elementos $\{\mathbf{x}, \mathbf{y}, \mathbf{z}, \ldots\}$ en el que se tienen definidas dos operaciones algebraicas, llamadas adición de vectores y multiplicación de vectores por escalares, recibe el nombre de **espacio vectorial** (o **espacio lineal**). Los elementos del espacio $S$ se llaman vectores. Si los escalares son números reales, el espacio se llama vectorial real; si los escalares comprenden los números complejos, el espacio se llama vectorial complejo. Las dos operaciones algebraicas indicadas son como sigue.

1. La **adición vectorial**, que asocia a cada par de elementos $\mathbf{x}, \mathbf{y} \in S$ un elemento único de $S$, conocido como su suma y denotado por $\mathbf{x} + \mathbf{y}$, debe satisfacer los axiomas siguientes:

1.1. La adición es conmutativa. Para dos elementos cualesquiera $\mathbf{x}, \mathbf{y} \in S$

$$\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$$

1.2. La adición en asociativa. Para tres elementos cualesquiera $\mathbf{x}, \mathbf{y}, \mathbf{z} \in S$

$$(\mathbf{x}+\mathbf{y})+\mathbf{z}=\mathbf{x}+(\mathbf{y}+\mathbf{z})$$

1.3. Existe un elemento único, llamado cero y denotado por $\mathbf{0}$, tal que para toda $\mathbf{x}$

$$\mathbf{x}+\mathbf{0}=\mathbf{x}$$

1.4. Para cada elemento $\mathbf{x} \in S$, existe un elemento único, llamado su negativo y denotado por $(-\mathbf{x})$, tal que

$$\mathbf{x}+(-\mathbf{x})=\mathbf{0}$$

**2**. La **multiplicación por escalares** asocia con cada escalar $\alpha$ (un número real o complejo, ya sea el caso) y con cada elemento $\mathbf{x} \in S$ un elemento único de $S$, llamado el producto de $\alpha$ y $\mathbf{x}$ y denotado por $\alpha\mathbf{x}$ (o $\mathbf{x}\alpha$). La multiplicación debe satisfacer los axiomas siguientes:

2.1. La multiplicación por escalares es asociativa. Para todos los escalares $\alpha$ y $\beta$ y cualquier elemento $\mathbf{x}$

$$(\alpha\beta)\,\mathbf{x}=\alpha\,(\beta\mathbf{x})$$

2.2. Primera ley distributiva. Para todos los escalares $\alpha$ y $\beta$ y cualquier elemento $\mathbf{x}$

$$(\alpha+\beta)\,\mathbf{x}=\alpha\mathbf{x}+\beta\mathbf{x}$$

2.3. Segunda ley distributiva. Para cualquier escalar $\alpha$ y todos los elementos $\mathbf{x}$ y $\mathbf{y}$

$$\alpha\,(\mathbf{x}+\mathbf{y})=\alpha\mathbf{x}+\alpha\mathbf{y}$$

2.4. Multiplicación unitaria. Para todo $\mathbf{x}$

$$1\cdot\mathbf{x}=\mathbf{x}$$

2.5. La multiplicación por cero

$$0\cdot\mathbf{x}=\mathbf{0}$$

Nótese que, de los axiomas 2.2 y 2.5, se tiene

$$\mathbf{x}+(-1)\,\mathbf{x}=(1-1)\,\mathbf{x}=\mathbf{0}$$

Por lo tanto, del axioma 1.4 se obtiene la igualdad

$$(-\mathbf{x})=(-1)\,\mathbf{x}$$

**Ejemplo 5.1.** *Los vectores del espacio físico tridimensional con la suma de vectores y con la multiplicación de vectores por números reales, definidas de manera tradicional, forman el espacio vectorial real.*

**Ejemplo 5.2.** *Todas las matrices de $m \times n$ con la suma de matrices y con la multiplicación de matrices por números complejos, definidas como es costumbre para las matrices, forman el espacio vectorial complejo; en este espacio las matrices son consideradas como vectores.*

**Ejemplo 5.3.** *Todas las funciones reales, continuas junto con sus primeras $k$ derivadas en el intervalo cerrado $[a,b]$, constituyen el espacio vectorial $C^{(k)}[a,b]$, si la suma de vectores y la multiplicación por escalares se definen de manera natural:*

$$(\mathbf{x}+\mathbf{y})(t) = \mathbf{x}(t) + \mathbf{y}(t), \quad (\alpha\mathbf{x})(t) = \alpha\mathbf{x}(t) \quad \textit{para } t \in [a,b]$$

*Cuando $k=0$, se tiene el espacio de funciones continuas $C[a,b]$.*

**Ejemplo 5.4.** *Todas aquellas funciones reales $y(t)$ en $C^{(k)}[a,b]$ que resuelven la ecuación diferencial lineal homogénea*

$$a_0(t)\frac{d^k y}{dt^k} + a_1(t)\frac{d^{k-1}y}{dt^{k-1}} + \cdots + a_n(t)\,y = 0$$

*forman un espacio vectorial real, en donde las operaciones se definen como en el caso de $C^{(k)}[a,b]$.*

**Ejemplo 5.5.** *Suponemos que $U$ y $V$ son dos espacios vectoriales reales o complejos. El **espacio producto** $U \times V$ es el espacio vectorial de pares ordenados $(\mathbf{u},\mathbf{v})$ con $\mathbf{u}$ en $U$ y con $\mathbf{v}$ en $V$, en donde*

$$(\mathbf{u},\mathbf{v}) + (\mathbf{u}',\mathbf{v}') = (\mathbf{u}+\mathbf{u}', \mathbf{v}+\mathbf{v}'), \quad \alpha(\mathbf{u},\mathbf{v}) = (\alpha\mathbf{u}, \alpha\mathbf{v})$$

*usando los mismos escalares que en $U$ y $V$.*

### 5.1.2. Dimensión y base del espacio, coordenadas

Un concepto de gran importancia para definir la estructura de un espacio vectorial es su dimensionalidad. Para tratar este concepto es necesario el aspecto de combinación lineal de vectores.

Una **combinación lineal** de los vectores $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in S$ es una suma de la forma

$$\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \cdots + \alpha_m\mathbf{x}_m = \sum_{i=1}^{m} \alpha_i\mathbf{x}_i$$

en donde $\alpha_i$ son escalares.

**Definición 5.1.** *Se dice que un conjunto de vectores $\{\mathbf{x}_i,\ i=1,\ldots,m\}$ es **linealmente dependiente** si y sólo si existe un conjunto de escalares $\{\alpha_i,\ i=1,\ldots,m\}$, tales que al menos uno de aquellos no es cero y se cumple*

$$\sum_{i=1}^{m} \alpha_i\mathbf{x}_i = \mathbf{0} \tag{5.1}$$

*Si no es así, se dice que el conjunto $\{\mathbf{x}_i,\ i=1,\ldots,m\}$ es **linealmente independiente**.*

Si los vectores $\{\mathbf{x}_i, i = 1, \ldots, m\}$ son linealmente dependientes, entonces uno de estos vectores puede representarse como combinación lineal de los demás. Por ejemplo, si $\alpha_1 \neq 0$, entonces

$$\mathbf{x}_1 = -\sum_{i=2}^{m} \frac{\alpha_i}{\alpha_1} \mathbf{x}_i$$

**Definición 5.2.** *Se dice que un espacio lineal $S$ es de* ***dimensión finita*** *y que un número $n$ es su* ***dimensión*** *si en $S$ existe un conjunto linealmente independiente de $n$ vectores y cualquier conjunto de $n+1$ vectores es linealmente dependiente. Si en $S$ se puede encontrar un sistema infinito de vectores linealmente independientes, entonces el espacio $S$ es de* ***dimensión infinita.***

**Definición 5.3.** *Sea $S$ un espacio $n$-dimensional. Entonces, cualquier conjunto linealmente independiente $B$ de $n$ vectores ordenados $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$, se llama la* ***base del espacio*** *$S$.*

Es fácil ver que cada espacio vectorial, en realidad, tiene un número infinito de bases. Pues, si $B = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ es una base en $S$, entonces cualquier conjunto de la forma $\{\alpha_1\mathbf{e}_1, \alpha_2\mathbf{e}_2, \ldots, \alpha_n\mathbf{e}_n\}$ también es una base en $S$. Sin embargo, cada base contiene exactamente $n$ vectores.

Si $S$ es $n$-dimensional, existe cuando menos una base $B = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ y cualquier vector $\mathbf{x} \in S$ se representa en la forma

$$\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i \tag{5.2}$$

Para comprobar la relación, notamos que $n+1$ vectores $\mathbf{x}, \mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ son linealmente dependientes, es decir,

$$\alpha_0 \mathbf{x} + \sum_{i=1}^{n} \alpha_i \mathbf{e}_i = 0$$

en donde al menos uno de los números $\alpha_0, \alpha_1, \ldots, \alpha_n$ no es cero. Es claro que $\alpha_0 \neq 0$, porque los vectores de la base $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ son linealmente independientes. Entonces, se tiene la ecuación (5.2), en donde $x_i = -\alpha_i/\alpha_0$ $(i = 1, 2, \ldots, n)$.

Los números $x_i$ $(i = 1, 2, \ldots, n)$ se llaman coordenadas del vector $\mathbf{x}$ en la base $B$. Las coordenadas $x_i$ $(i = 1, 2, \ldots, n)$ son unívocamente determinadas por el vector $\mathbf{x}$ y la base $B$, es decir, el desarrollo (5.2) es único. Para comprobar el hecho, suponemos que existe alguna otra representación de $\mathbf{x}$,

$$\mathbf{x} = \sum_{i=1}^{n} x_i' \mathbf{e}_i$$

Restando ésta de (5.2), se tiene

$$\sum_{i=1}^{n} (x_i - x_i')\, \mathbf{e}_i = \mathbf{0}$$

Pues, $x_i - x'_i = 0\ (i = 1, 2, \ldots, n)$, de donde se sigue la unicidad de coordenadas de cada vector en $S$ con respecto a una base dada $B$.

Sea $S$ un espacio lineal $n$-dimensional y $B = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ una base en $S$. Si

$$\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i \quad \text{y} \quad \mathbf{y} = \sum_{i=1}^{n} y_i \mathbf{e}_i$$

son dos vectores cualesquiera en $S$, entonces la adición de vectores y la multiplicación de vectores por escalares se representa en coordenadas de la siguiente manera:

$$\mathbf{x} + \mathbf{y} = \sum_{i=1}^{n} (x_i + y_i)\, \mathbf{e}_i, \quad \alpha \mathbf{x} = \sum_{i=1}^{n} \alpha x_i \mathbf{e}_i \tag{5.3}$$

es decir, la coordenada $i$-ésima de la suma es la suma de las coordenadas $i$-ésimas y la coordenada $i$-ésima del producto es el producto de un escalar por la coordenada correspondiente del vector.

### 5.1.3. Subespacios

Un subconjunto no vacío $\tilde{S} \subseteq S$ de un espacio vectorial $S$ recibe el nombre de **subespacio** de éste, si el propio $\tilde{S}$ es un espacio vectorial con respecto a las operaciones algebraicas definidas en $S$. Es fácil demostrar que un subconjunto no vacío $\tilde{S}$ de un espacio vectorial $S$ es un subespacio de $S$, si y sólo si para $\mathbf{x}$ y $\mathbf{y}$ cualesquiera en $\tilde{S}$ y cualquier escalar $\alpha$,

$$(\mathbf{x} + \mathbf{y}) \in \tilde{S} \quad \text{y} \quad \alpha \mathbf{x} \in \tilde{S}$$

Es decir, $\tilde{S}$ es cerrado con respecto tanto a la suma de vectores como a la multiplicación de vectores por escalares.

**Ejemplo 5.6.** *Cualquier plano bidimensional es un subespacio en el espacio físico tridimensional.*

**Ejemplo 5.7.** *El conjunto $P^k$ de todos los polinomios de grado menor o igual a $(k-1)$ con coeficientes reales es un subespacio de $C^{(k-1)}[a,b]$.*

**Ejemplo 5.8.** *El espacio $C^{(k)}[a,b]$ de funciones reales, continuas junto con sus primeras $k$ derivadas en el intervalo cerrado $[a,b]$, constituyen un subespacio de $C[a,b]$.*

### 5.1.4. Descomposición de un espacio en la suma de subespacios

**Definición 5.4.** *Si $S'$ y $S''$ son dos subespacios de $S$ y, además, se conoce que*

1) $S'$ y $S''$ no tienen los vectores en común, excepto el vector $\mathbf{0}$, y
2) cualquier vector $\mathbf{x}$ en $S$ se representa como la suma de dos vectores

$$\mathbf{x} = \mathbf{x}' + \mathbf{x}'' \quad (\mathbf{x}' \in S', \mathbf{x}'' \in S'') \tag{5.4}$$

se dice que el espacio $S$ se descompone en la **suma de dos subespacios** $S'$ y $S''$ y se escribe

$$S = S' + S'' \tag{5.5}$$

Nótese que la primera condición de la definición implica la unicidad de la descomposición (5.4). En realidad, si para un vector $\mathbf{x}$ se tenían dos distintas representaciones, en la forma de suma de dos vectores en $S'$ y $S''$, digamos (5.4) y

$$\mathbf{x} = \tilde{\mathbf{x}}' + \tilde{\mathbf{x}}'' \quad (\tilde{\mathbf{x}}' \in S',\ \tilde{\mathbf{x}}'' \in S'')$$

entonces, restando una de otra, se tiene la igualdad

$$\mathbf{x}' - \tilde{\mathbf{x}}' = \tilde{\mathbf{x}}'' - \mathbf{x}''$$

de dos vectores distintos de cero, $\mathbf{x}' - \tilde{\mathbf{x}}' \in S'$ y $\tilde{\mathbf{x}}'' - \mathbf{x}'' \in S''$. Esta igualdad contradice a la primera condición, comprobando la unicidad de la representación (5.4).

Es fácil extender la definición de la descomposición de un espacio en la suma de varios subespacios.

Sea $S = S' + S''$ y sean $\{\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_k\}$ y $\{\mathbf{e}''_1, \mathbf{e}''_2, \ldots, \mathbf{e}''_m\}$ bases de $S'$ y $S''$, respectivamente. Entonces, es fácil demostrar (se queda como tarea para el estudiante) que dichos $k+m$ vectores son linealmente independientes y constituyen una base en $S$, es decir, la base del espacio $S$ se compone de las bases de los subespacios sumados. En particular, se concluye que $\dim(S) = \dim(S') + \dim(S'')$. Al escoger una base $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ en un espacio $S$, en realidad se fija cierta descomposición de todo el espacio $S$ en la suma de $n$ subespacios unidimensionales.

**Problema.** *Sea $\tilde{S}$ un subespacio $m$-dimensional en un espacio $n$-dimensional $S$. Demuestre que $m \leq n$ y, cuando $m = n$, entonces $\tilde{S} \equiv S$.*

## 5.2. Transformaciones lineales y matrices

### 5.2.1. Correspondencia entre matrices y operadores lineales

La importancia del concepto de espacios lineales viene no sólo de la capacidad de representar con vectores muchos objetos de interés, sino también de representar muchas de las operaciones o transformaciones sobre esos objetos en forma de matrices. Podemos mencionar modelos de sistemas o procesos en donde los datos de entrada se transforman en los de salida. Además, en las Matemáticas a menudo se encuentra la situación del mapeo de un espacio (o dominio) a otro y/o el cambio de una base y sistema de coordenadas.

**Definición 5.5.** *Sean $X$ y $Y$ espacios vectoriales, ambos reales o complejos. Una **transformación lineal** $\hat{A}$ de $X$ a $Y$ es una correspondencia que asigna*

*a cada vector* $\mathbf{x}$ *de* $X$ *un vector* $\hat{A}\mathbf{x}$ *de* $Y$ *de tal manera que*

$$\hat{A}(\mathbf{x}_1 + \mathbf{x}_2) = \hat{A}\mathbf{x}_1 + \hat{A}\mathbf{x}_2 \tag{5.6}$$

$$\hat{A}(\alpha\mathbf{x}) = \alpha\hat{A}\mathbf{x} \tag{5.7}$$

*para cualquier par de vectores* $\mathbf{x}_1$ *y* $\mathbf{x}_2$ *de* $X$ *y cualquier escalar* $\alpha$ *real o complejo, según sea el caso. Se dice también que dicha transformación define el* ***operador lineal*** $\hat{A}$ *con el dominio de definición* $X$ *y el contradominio* $\hat{A}(X) \subseteq Y$.

Una clase especial de transformaciones lineales son las de un espacio a sí mismo, de $X$ a $X$.

**Ejemplo 5.9.** *Sea* $X$ *un espacio físico tridimensional y* $Y$ *un plano en éste. La proyección de* $X$ *sobre el plano* $Y$ *es una transformación lineal.*

**Ejemplo 5.10.** *Sea* $X = C^{(2)}(0,1)$ *el espacio de funciones diferenciables dos veces. Asignamos a cada función* $x(t)$ *de* $X$ *una función mediante la relación*

$$y(t) \equiv \hat{A}(x(t)) = \frac{d}{dt}\left((1+t)\frac{dx}{dt}\right) - (2+e^t)\,x(t)$$

*Es fácil ver que* $\hat{A}$ *es una transformación (un operador) lineal de* $X$ *a* $Y = C(0,1)$.

El hecho fundamental con respecto a las transformaciones lineales definidas entre espacios vectoriales de dimensión finita es que existe una relación estrecha entre las transformadas y matrices que se demuestra a continuación.

Sean $X$ un espacio vectorial $n$-dimensional con una base $B = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ e $Y$ otro espacio vectorial $m$-dimensional que tiene una base $\tilde{B} = \{\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \ldots, \tilde{\mathbf{e}}_m\}$. Sea $\hat{A}$ cualquier operador (transformación) lineal de $X$ a $Y$. Aplicando este operador a los vectores de la base $B$, se obtienen

$$\mathbf{a}_j = \hat{A}\mathbf{e}_j, \quad j = 1, 2, \ldots, n$$

Cada vector $\mathbf{a}_i$ del espacio $Y$ puede desarrollarse en la base $\tilde{B}$

$$\mathbf{a}_j = \hat{A}\mathbf{e}_j = \sum_{i=1}^{m} a_{ij}\tilde{\mathbf{e}}_i, \quad j = 1, 2, \ldots, n \tag{5.8}$$

en donde $a_{ij}$ es la $i$-ésima coordenada del vector $\mathbf{a}_j$ en la base $\tilde{B}$. Sea $\mathbf{y} = \hat{A}\mathbf{x}$, en donde $\mathbf{x}$ es cualquier vector de $X$ y sean

$$\mathbf{x} = \sum_{j=1}^{n} x_j\mathbf{e}_j, \quad \mathbf{y} = \sum_{i=1}^{m} y_i\tilde{\mathbf{e}}_i$$

desarrollos de dichos vectores en la base $B$ y $\tilde{B}$, respectivamente. A continuación,

$$\begin{aligned}\sum_{i=1}^{m} y_i \tilde{\mathbf{e}}_i &= \mathbf{y} = \hat{A}\mathbf{x} \\ &= \hat{A}\left(\sum_{j=1}^{n} x_j \mathbf{e}_j\right) = \sum_{j=1}^{n} x_j \hat{A}\mathbf{e}_j = \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij}\tilde{\mathbf{e}}_i \\ &= \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}x_j\right)\tilde{\mathbf{e}}_i\end{aligned}$$

de donde

$$y_i = \sum_{j=1}^{n} a_{ij}x_j, \quad (i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n) \tag{5.9}$$

es decir, las coordenadas del vector transformado $\mathbf{y} = \hat{A}\mathbf{x}$ se expresan a través de las coordenadas del vector original $\mathbf{x}$ mediante una matriz $A = [a_{ij}]$. Escribiendo como $\vec{x} = (x_1, x_2, \ldots, x_n)$ y $\vec{y} = (y_1, y_2, \ldots, y_m)$ vectores columnas de coordenadas de los vectores $\mathbf{x}$ y $\mathbf{y}$, respectivamente, la relación (5.9) se escribe en la forma matricial

$$\vec{y} = A\vec{x} \tag{5.10}$$

que corresponde a la transformada $\mathbf{y} = \hat{A}\mathbf{x}$.

La relación (5.9) demuestra la propiedad fundamental de las transformaciones lineales: dadas las bases en un espacio $n$-dimensional $X$ y $m$-dimensional $Y$, a cada operador (transformación) lineal $\hat{A}$, que mapea $X$ en $Y$, le corresponde una matriz rectangular $A = [a_{ij}]$ de $m \times n$, y viceversa, a cada matriz $A = [a_{ij}]$ de $m \times n$ le corresponde un operador lineal que mapea $X$ en $Y$. La matriz $A = [a_{ij}]$ se define mediante la relación (5.8).

### 5.2.2. Espacio imagen y espacio nulo

La transformación $\mathbf{y} = \hat{A}\mathbf{x}$, cuando $\mathbf{y}$ está dado, puede considerarse como una ecuación con respecto a la incógnita $\mathbf{x}$. La $\mathbf{y} = \hat{A}\mathbf{x}$, una vez representada en la forma matricial (5.10), puede tratarse como un sistema de ecuaciones lineales. Por lo tanto, la resolución de $\mathbf{y} = \hat{A}\mathbf{x}$ está relacionada con la del sistema de ecuaciones lineales $\vec{y} = A\vec{x}$, cuya resolución se ha tratado en el capítulo 2. Se sabe que la solución de un sistema de ecuaciones lineales $\vec{y} = A\vec{x}$ no es siempre única (en caso de existir una solución); para un vector dado $\vec{y}$ pueden existir varios $\vec{x}$ que satisfacen la ecuación. Para entender esta relación más a detalle en términos de espacios vectoriales $X$ y $Y$, son útiles los conceptos de espacio imagen y espacio nulo.

Considérese el conjunto $\hat{A}(X) \subseteq Y$ de todos los vectores $\mathbf{y} \in Y$ tales que $\mathbf{y} = \hat{A}\mathbf{x}$ para algún $\mathbf{x}$ en $X$. Es fácil comprobar que $\hat{A}(X)$ es un subespacio de $Y$. El subespacio $\hat{A}(X) \subseteq Y$ recibe el nombre de **espacio imagen** de la

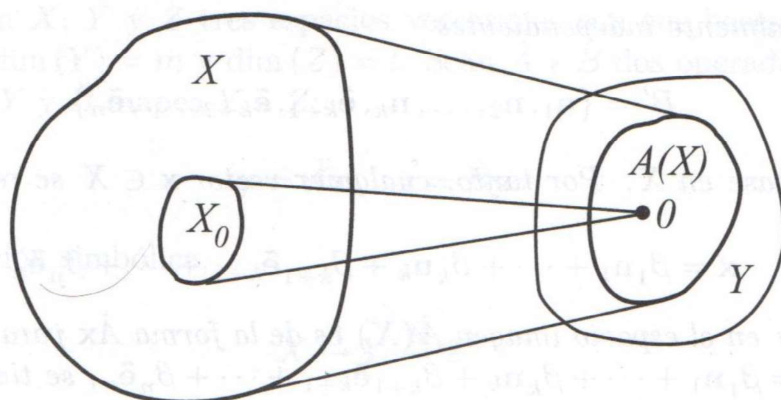



Figura 5.1. Espacio imagen y espacio nulo

transformación $\hat{A}$ de $X$ a $Y$. Con esta definición, es fácil ver que la ecuación $\mathbf{y} = \hat{A}\mathbf{x}$ se puede resolver con respecto a la incógnita $\mathbf{x}$ si y sólo si $\mathbf{y}$ está en el espacio imagen $\hat{A}(X)$ de $\hat{A}$; dicha ecuación se puede también resolver para cualquier $\mathbf{y}$ si el espacio imagen es igual al contradominio de $\hat{A}$, $\hat{A}(X) = Y$.

Para contestar la pregunta: ¿cuándo la solución de $\mathbf{y} = \hat{A}\mathbf{x}$ es única?, suponemos que $\mathbf{x}_1$ y $\mathbf{x}_2$ son dos soluciones distintas, $\mathbf{y} = \hat{A}\mathbf{x}_1$ y $\mathbf{y} = \hat{A}\mathbf{x}_2$. Por lo tanto,

$$\mathbf{0} = \mathbf{y} - \mathbf{y} = \hat{A}\mathbf{x}_1 - \hat{A}\mathbf{x}_2 = \hat{A}(\mathbf{x}_1 - \mathbf{x}_2)$$

así que $(\mathbf{x}_1 - \mathbf{x}_2)$ es una solución de la ecuación homogénea $\hat{A}\mathbf{x} = \mathbf{0}$. De nuevo, es fácil ver que el conjunto $X_0$ de todas las soluciones de $\hat{A}\mathbf{x} = \mathbf{0}$ es un subespacio de $X$; este subespacio se llama **espacio nulo** (o núcleo) de la transformación lineal $\hat{A}$ de $X$ a $Y$. Con base en la consideración anterior, concluimos que la ecuación $\mathbf{y} = \hat{A}\mathbf{x}$ tiene una solución única $\mathbf{x}$, si y sólo si el espacio nulo $X_0$ de $\hat{A}$ consiste en un solo elemento $\mathbf{0}$.

Las relaciones entre el dominio $X$, el contradominio $Y$, el espacio imagen $\hat{A}(X)$ y el espacio nulo $X_0$ para una transformación lineal $\hat{A}$ se ilustran esquemáticamente en la figura 5.1.

El esquema anterior sugiere el "encogimiento" del espacio nulo $X_0 \subset X$ hasta el vector $\mathbf{0}$ en $Y$ y el "encogimiento" similar de $X$ hasta el espacio imagen $\hat{A}(X) \subseteq Y$. El siguiente teorema demuestra que esto es correcto en términos de dimensiones.

**Teorema 5.1.** *Sea $\hat{A}$ una transformación lineal de un espacio de dimensión finita $X$ a un espacio $Y$. Entonces,*

$$\dim(X) = \dim(X_0) + \dim\left(\hat{A}(X)\right) \tag{5.11}$$

*es decir, la dimensión de dominio $X$ es igual a la suma de la dimensión del espacio nulo $X_0$ y la dimensión del espacio imagen $\hat{A}(X) \subseteq Y$.*

**Demostración.** *Sea $\dim(X) = n$ y $\dim(X_0) = k \leq n$. Sea $\{\mathbf{n}_1, \mathbf{n}_2, \ldots, \mathbf{n}_k\}$ una base para $X_0$. La completamos con $n - k$ vectores linealmente independientes $\{\tilde{\mathbf{e}}_{k+1}, \tilde{\mathbf{e}}_{k+2}, \ldots, \tilde{\mathbf{e}}_n\}$, que no pertenecen a $X_0$, hasta tener un conjunto de $n$*

*vectores linealmente independientes*

$$B' = \{\mathbf{n}_1, \mathbf{n}_2, \ldots, \mathbf{n}_k, \tilde{\mathbf{e}}_{k+1}, \tilde{\mathbf{e}}_{k+2}, \ldots, \tilde{\mathbf{e}}_n\}$$

*que es una base en $X$. Por tanto, cualquier vector $\mathbf{x} \in X$ se representa en la forma*

$$\mathbf{x} = \beta_1 \mathbf{n}_1 + \cdots + \beta_k \mathbf{n}_k + \beta_{k+1} \tilde{\mathbf{e}}_{k+1} + \cdots + \beta_n \tilde{\mathbf{e}}_n$$

*Como toda $\mathbf{y}$ en el espacio imagen $\hat{A}(X)$ es de la forma $\hat{A}\mathbf{x}$ para alguna $\mathbf{x} \in X$, y ya que $\mathbf{x} = \beta_1 \mathbf{n}_1 + \cdots + \beta_k \mathbf{n}_k + \beta_{k+1} \tilde{\mathbf{e}}_{k+1} + \cdots + \beta_n \tilde{\mathbf{e}}_n$, se tiene que*

$$\mathbf{y} = \hat{A}\mathbf{x} = \beta_{k+1} \hat{A}\tilde{\mathbf{e}}_{k+1} + \cdots + \beta_n \hat{A}\tilde{\mathbf{e}}_n$$

*en donde se aplicó $\hat{A}\mathbf{n}_i = 0$, pues los vectores $\hat{A}\tilde{\mathbf{e}}_{k+1}, \ldots, \hat{A}\tilde{\mathbf{e}}_n$ generan el espacio imagen $\hat{A}(X)$. Además, estos vectores son linealmente independientes. Si no fuera así, se tendría $\mathbf{0} = \tilde{\beta}_{k+1} \hat{A}\tilde{\mathbf{e}}_{k+1} + \cdots + \tilde{\beta}_n \hat{A}\tilde{\mathbf{e}}_n = \hat{A}\left(\tilde{\beta}_{k+1} \tilde{\mathbf{e}}_{k+1} + \cdots + \tilde{\beta}_n \tilde{\mathbf{e}}_n\right)$, lo que no es posible porque el vector $\tilde{\beta}_{k+1} \tilde{\mathbf{e}}_{k+1} + \cdots + \tilde{\beta}_n \tilde{\mathbf{e}}_n$ no pertenece a $X_0$. Entonces $\left\{\hat{A}\tilde{\mathbf{e}}_{k+1}, \ldots, \hat{A}\tilde{\mathbf{e}}_n\right\}$ es una base de $\hat{A}(X)$, $\dim\left(\hat{A}(X)\right) = n - k$, y esto completa la demostración de (5.11).* ■

### 5.2.3. Suma y producto de operadores lineales

Sean $\hat{A}$ y $\hat{B}$ dos operadores lineales que mapean $X$ a $Y$, $\dim(X) = n$ y $\dim(Y) = m$. Sean $B = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ y $\tilde{B} = \{\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \ldots, \tilde{\mathbf{e}}_m\}$ bases de $X$ a $Y$, respectivamente, y

$$A = [a_{ij}] \quad \text{y} \quad B = [b_{ij}] \quad (i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n)$$

dos matrices correspondientes.

**Definición 5.6.** *La suma de operadores lineales $\hat{A}$ y $\hat{B}$ es el operador $\hat{C} = \hat{A} + \hat{B}$ determinado mediante la igualdad*

$$\hat{C}\mathbf{x} = \hat{A}\mathbf{x} + \hat{B}\mathbf{x}, \quad \textit{para toda } \mathbf{x} \in X \tag{5.12}$$

Con base en esta definición, es fácil ver que la suma $\hat{C} = \hat{A} + \hat{B}$ es también un operador lineal y $\hat{C} = \hat{A} + \hat{B} = \hat{B} + \hat{A}$. Además,

$$\hat{C}\mathbf{e}_j = \hat{A}\mathbf{e}_j + \hat{B}\mathbf{e}_j = \sum_{i=1}^{m} (a_{ij} + b_{ij})\, \tilde{\mathbf{e}}_i$$

Así que al operador $\hat{C}$ le corresponde la matriz

$$C = A + B \tag{5.13}$$

donde $C = [c_{ij}]$, $c_{ij} = a_{ij} + b_{ij}$ $(i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n)$.

Ahora, sean $X$, $Y$ y $Z$ tres espacios vectoriales con sus bases respectivas; $\dim(X) = n$, $\dim(Y) = m$ y $\dim(Z) = l$. Sean $\hat{A}$ y $\hat{B}$ dos operadores lineales, $\hat{A}$ mapea $X$ a $Y$ y $\hat{B}$ mapea $Y$ a $Z$:

$$\mathbf{y} = \hat{A}\mathbf{x}, \quad \mathbf{z} = \hat{B}\mathbf{y}$$

En representación simbólica,

$$X \xrightarrow{\hat{A}} Y \xrightarrow{\hat{B}} Z$$

**Definición 5.7.** *El producto de dos operadores $\hat{B}$ y $\hat{A}$ es el operador $\hat{C}$ tal que, para cualquier $\mathbf{x} \in X$,*

$$\hat{C}\mathbf{x} = \hat{B}\left(\hat{A}(\mathbf{x})\right) \equiv \hat{B}\hat{A}\mathbf{x} \tag{5.14}$$

*El operador $\hat{C}$ mapea $X$ a $Z$,*

$$X \overset{\hat{C}=\hat{B}\hat{A}}{\rightarrow} Z$$

*dado que $\mathbf{z} = \hat{B}(\mathbf{y}) = \hat{B}\left(\hat{A}(\mathbf{x})\right)$.*

El producto $\hat{B}\hat{A}$ es un operador lineal, si $\hat{A}$ y $\hat{B}$ son lineales:

$$\hat{B}\hat{A}(\alpha\mathbf{x}) = \hat{B}\left(\alpha\left(\hat{A}\mathbf{x}\right)\right) = \alpha\hat{B}\hat{A}\mathbf{x}$$

$$\hat{B}\hat{A}(\mathbf{x}_1 + \mathbf{x}_2) = \hat{B}\left(\hat{A}\mathbf{x}_1 + \hat{A}\mathbf{x}_2\right) = \hat{B}\hat{A}\mathbf{x}_1 + \hat{B}\hat{A}\mathbf{x}_2$$

Sean $B = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$, $\tilde{B} = \{\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \ldots, \tilde{\mathbf{e}}_m\}$ y $B' = \{\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_l\}$ bases cualesquiera en $X$, $Y$ y $Z$, respectivamente, y sean

$$\begin{aligned} A &= [a_{ij}], \quad \hat{A}\mathbf{e}_j = \sum_{i=1}^{m} a_{ij}\tilde{\mathbf{e}}_i, \quad (j = 1, 2, \ldots, n) \\ B &= [b_{ki}], \quad \hat{B}\tilde{\mathbf{e}}_i = \sum_{k=1}^{l} b_{ki}\mathbf{e}'_k, \quad (i = 1, 2, \ldots, m) \\ C &= [c_{kj}], \quad \hat{C}\mathbf{e}_j = \sum_{k=1}^{l} c_{kj}\mathbf{e}'_k, \quad (j = 1, 2, \ldots, n) \end{aligned}$$

matrices correspondientes a los operadores $\hat{A}$, $\hat{B}$ y $\hat{C}$ en las bases escogidas. Según (5.9), a las igualdades vectoriales

$$\mathbf{y} = \hat{A}\mathbf{x}, \quad \mathbf{z} = \hat{B}\mathbf{y}, \quad \mathbf{z} = \hat{C}\mathbf{x}$$

les corresponden las igualdades matriciales

$$\vec{y} = A\vec{x}, \quad \vec{z} = B\vec{y}, \quad \vec{z} = C\vec{x}$$

en donde $\vec{x}$, $\vec{y}$ y $\vec{z}$ son vectores columnas de coordenadas de $\mathbf{x}$, $\mathbf{y}$ y $\mathbf{z}$. Comparando $\vec{z} = C\vec{x}$ y $\vec{z} = B\vec{y} = B(A\vec{x})$, es decir,

$$z_k = \sum_{i=1}^{m} b_{ki}y_i = \sum_{i=1}^{m} b_{ki} \sum_{j=1}^{n} a_{ij}x_j = \sum_{j=1}^{n} \left( \sum_{i=1}^{m} b_{ki}a_{ij} \right) x_j$$

y

$$z_k = \sum_{j=1}^{n} c_{kj}x_j$$

se tiene

$$c_{kj} = \sum_{i=1}^{m} b_{ki}a_{ij} \quad \text{o} \quad C = BA \tag{5.15}$$

Entonces, al producto $\hat{C} = \hat{B}\hat{A}$ de operadores $\hat{B}$ y $\hat{A}$ le corresponde la matriz $C = [c_{kj}]$ $(k = 1, 2, \ldots, l;\ j = 1, 2, \ldots, n)$, que es igual al producto de matrices $B$ y $A$, ecuación (5.15).

Es fácil ver que al operador $\hat{C} = \alpha\hat{A}$ le corresponde la matriz $C = \alpha A$, en donde $\alpha$ es un escalar.

Resumiendo los resultados de esta sección, vemos que a la suma $\hat{A} + \hat{B}$, al producto $\hat{C} = \hat{B}\hat{A}$ y $\alpha\hat{A}$ de operadores lineales les corresponden unívocamente las operaciones matriciales $A + B$, $C = BA$ y $\alpha A$.

### 5.2.4. Conmutatividad de transformaciones lineales

Dados dos operadores (transformaciones) lineales $\hat{A}$ y $\hat{B}$ que transforman $X$ a $X$, se tienen $\hat{C} = \hat{B}\hat{A}$ y $\hat{D} = \hat{A}\hat{B}$. Por lo general, $\hat{B}\hat{A} \neq \hat{A}\hat{B}$, es decir, existe cuando menos un vector $\mathbf{x} \in X$ tal que $\hat{B}\hat{A}\mathbf{x} \neq \hat{A}\hat{B}\mathbf{x}$.

**Definición 5.8.** *Si $\hat{B}\hat{A} = \hat{A}\hat{B}$ para toda $\mathbf{x} \in X$, se dice que los operadores $\hat{A}$ y $\hat{B}$ conmutan o son conmutativos.*

Se puede demostrar que si los operadores conmutan, entonces existe una base $B$ en $X$, en donde las matrices de representación de los operadores son diagonales simultáneamente, y viceversa (véase, por ejemplo, [8]).

Para cada espacio vectorial $X$, se define el **operador de identidad** (unitario) $\hat{I}$ que mapea $X$ a $X$ de la siguiente manera:

$$\hat{I}\mathbf{x} = \mathbf{x} \tag{5.16}$$

para toda $\mathbf{x} \in X$. Si $\hat{A}$ es cualquier operador lineal que transforma $X$ a $Y$, entonces, para cualquier $\mathbf{x}$,

$$\hat{A}\left(\hat{I}\mathbf{x}\right) = \hat{A}\mathbf{x} \Rightarrow \hat{A}\hat{I} = \hat{A} \tag{5.17}$$

Para cualquier base en $X$, la matriz de representación de $\hat{I}$ es

$$I = [\delta_{ij}] \quad (i, j = 1, 2, \ldots, \dim(X)) \tag{5.18}$$




donde $\delta_{ij}$ es la delta de Kronecker.

Si $\hat{I}$ es el operador unitario en $Y$ y $\mathbf{y} = \hat{A}\mathbf{x}$, entonces,

$$\hat{A}\mathbf{x} = \mathbf{y} = \hat{I}\mathbf{y} = \hat{I}\hat{A}\mathbf{x} \Rightarrow \hat{A} = \hat{I}\hat{A} \tag{5.19}$$

Para cada espacio vectorial $X$ se puede definir el **operador nulo** $\hat{0}$ que transforma $X$ en $\mathbf{0}$:

$$\hat{0}\mathbf{x} = \mathbf{0}$$

para toda $\mathbf{x} \in X$, es decir, $\hat{A}\hat{0} = \hat{0}$.

### 5.2.5. Inversas de transformaciones lineales

Dada una transformación lineal $\hat{A}$ de $X$ a $Y$, $\mathbf{y} = \hat{A}\mathbf{x}$, el problema inverso consiste en determinar $\mathbf{x}$, conocida $\mathbf{y}$. La resolución del problema propuesto la hemos tratado brevemente en la sección 5.2.2. El resumen es que la ecuación $\mathbf{y} = \hat{A}\mathbf{x}$ tiene una solución única $\mathbf{x} \in X$ para cada $\mathbf{y} \in \hat{A}(X)$ si y sólo si el espacio nulo $X_0$, $\hat{A}(X_0) = \mathbf{0}$, consiste en un solo elemento, $X_0 = \{\mathbf{0}\}$. En este caso, $\dim(X) = \dim\left(\hat{A}(X)\right)$. La correspondencia de $\mathbf{y}$ a la solución $\mathbf{x}$ se llama la **transformación inversa** de $\hat{A}$ y se denota como $\mathbf{x} = \hat{A}^{-1}\mathbf{y}$, de modo que

$$\mathbf{y} = \hat{A}\mathbf{x} = \hat{A}\left(\hat{A}^{-1}\mathbf{y}\right) \quad \text{y} \quad \hat{A}^{-1}\left(\hat{A}\mathbf{x}\right) = \hat{A}^{-1}\mathbf{y} = \mathbf{x} \tag{5.20}$$

para toda $\mathbf{y} \in \hat{A}(X)$ y toda $\mathbf{x} \in X$. Se ve que $\hat{A}^{-1}$ transforma $Y$ en $X$. Utilizando el operador de identidad, las igualdades anteriores se escriben como

$$\hat{A}\hat{A}^{-1} = \hat{I}_Y \quad \text{y} \quad \hat{A}^{-1}\hat{A} = \hat{I}_X \tag{5.21}$$

En el caso particular cuando un operador $\hat{A}$ mapea $X$ en $X$ (se dice que es un mapeo de $X$ en sí mismo), $\hat{A}(X) = X$, se tiene

$$\hat{A}\hat{A}^{-1} = \hat{A}^{-1}\hat{A} = \hat{I} \tag{5.22}$$

Comprobamos que el operador inverso $\hat{A}^{-1}$ es lineal. Sean $\mathbf{x}_1 = \hat{A}^{-1}\mathbf{y}_1$ y $\mathbf{x}_2 = \hat{A}^{-1}\mathbf{y}_2$ soluciones únicas de $\mathbf{y}_1 = \hat{A}\mathbf{x}_1$ y $\mathbf{y}_2 = \hat{A}\mathbf{x}_2$, respectivamente. Debido a que $\hat{A}$ es lineal,

$$\hat{A}(\mathbf{x}_1 + \mathbf{x}_2) = \hat{A}\mathbf{x}_1 + \hat{A}\mathbf{x}_2 = \mathbf{y}_1 + \mathbf{y}_2$$

Entonces, con la definición (5.20),

$$\hat{A}^{-1}(\mathbf{y}_1 + \mathbf{y}_2) = \hat{A}^{-1}\left(\hat{A}(\mathbf{x}_1 + \mathbf{x}_2)\right) = \mathbf{x}_1 + \mathbf{x}_2 = \hat{A}^{-1}\mathbf{y}_1 + \hat{A}^{-1}\mathbf{y}_2$$

Además,

$$\hat{A}(\alpha\mathbf{x}_1) = \alpha\hat{A}\mathbf{x}_1 = \alpha\mathbf{y}_1$$

y, por lo tanto,

$$\hat{A}^{-1}(\alpha\mathbf{y}_1) = \hat{A}^{-1}\left(\hat{A}(\alpha\mathbf{x}_1)\right) = \alpha\mathbf{x}_1 = \alpha\hat{A}^{-1}\mathbf{y}_1$$

Queda completamente comprobada la linealidad de la transformación inversa $\hat{A}^{-1}$.

Como hemos visto, la resolución de un problema inverso es equivalente a encontrar la inversa de la transformada $\hat{A}$, $\mathbf{x} = \hat{A}^{-1}\mathbf{y}$. En la práctica, de manera explícita se resuelven ecuaciones matriciales, sistemas de ecuaciones lineales, con una representación matricial particular $A$ del operador $\hat{A}$. En este caso, al operador $\hat{A}^{-1}$ le corresponde la matriz $A^{-1}$ (véase el capítulo 2 y cursos de álgebra lineal estándares). Sin embargo, el concepto del operador inverso (la inversa de una transformada) es de gran utilidad para las demostraciones y cálculos fundamentales.

## 5.3. Cambio de base

### 5.3.1. Transformación de coordenadas

Considérense las dos bases $B = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ y $B' = \{\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n\}$ en un espacio lineal $n$-dimensional $X$. La posición mutua de dos bases se define mediante las coordenadas de los vectores de una base con respecto a la otra. Sea

$$\mathbf{e}_i = \sum_{j=1}^{n} T_{ji}\mathbf{e}'_j, \quad (i = 1, 2, \ldots, n) \tag{5.23}$$

en donde $T_{ji}$ $(i, j = 1, 2, \ldots, n)$ son coordenadas del vector $\mathbf{e}_i$ con respecto a la base $B'$.

Establecemos la relación entre las coordenadas de un vector en diferentes bases. Si un vector $\mathbf{x} \in X$ tiene coordenadas $x_1, x_2, \ldots, x_n$ y $x'_1, x'_2, \ldots, x'_n$ en $B$ y $B'$, respectivamente, entonces

$$\mathbf{x} = \sum_{i=1}^{n} x'_i\mathbf{e}'_i = \sum_{i=1}^{n} x_i\mathbf{e}_i = \sum_{i=1}^{n} x_i \sum_{j=1}^{n} T_{ji}\mathbf{e}'_j = \sum_{j=1}^{n}\left(\sum_{i=1}^{n} T_{ji}x_i\right)\mathbf{e}'_j$$

De donde, tomando en cuenta que las coordenadas de un vector se determinan unívocamente por la base, se obtiene

$$x'_j = \sum_{i=1}^{n} T_{ji}x_i, \quad (j = 1, 2, \ldots, n) \tag{5.24}$$

que en forma matricial es

$$\vec{x}' = T\vec{x} \tag{5.25}$$

en donde $\vec{x}$ y $\vec{x}'$ son vectores columnas de coordenadas de $\mathbf{x}$ en $B$ y $B'$, respectivamente, y $T = [T_{ji}]$ es la matriz de coordenadas de los vectores de la base $B$ con respecto a la base $B'$, ecuación (5.23).

La matriz $T$ se llama matriz de transformación de coordenadas y es no singular

$$\det T \neq 0 \tag{5.26}$$

lo que se comprueba a continuación.

Por analogía con (5.23), se tiene

$$\mathbf{e}'_i = \sum_{j=1}^{n} \left(T^{-1}\right)_{ji} \mathbf{e}_j, \quad (i = 1, 2, \ldots, n) \tag{5.27}$$

en donde $T^{-1}$ es la matriz de coordenadas de vectores de $B'$ con respecto a la base $B$. Para un vector $\mathbf{x}$, se tiene

$$\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i = \sum_{i=1}^{n} x'_i \mathbf{e}'_i = \sum_{i=1}^{n} x'_i \sum_{j=1}^{n} \left(T^{-1}\right)_{ji} \mathbf{e}_j = \sum_{j=1}^{n} \left( \sum_{i=1}^{n} \left(T^{-1}\right)_{ji} x'_i \right) \mathbf{e}_j$$

Entonces,

$$x_j = \sum_{i=1}^{n} \left(T^{-1}\right)_{ji} x'_i, \quad (j = 1, 2, \ldots, n) \tag{5.28}$$

que en forma matricial es

$$\vec{x} = T^{-1} \vec{x}' \tag{5.29}$$

en donde $\vec{x}$ y $\vec{x}'$ son vectores columnas de coordenadas de $\mathbf{x}$ en $B$ y $B'$, respectivamente, y $T^{-1} = \left[\left(T^{-1}\right)_{ji}\right]$ es la matriz de coordenadas de vectores de la base $B'$ con respecto a la base $B$, ecuación (5.27).

La comparación de (5.29) con (5.25) nos sugiere que $T^{-1}$ es la inversa de $T$. Para comprobarlo, sustituimos (5.27) en (5.23)

$$\mathbf{e}_i = \sum_{j=1}^{n} T_{ji} \sum_{k=1}^{n} \left(T^{-1}\right)_{kj} \mathbf{e}_k = \sum_{k=1}^{n} \left( \sum_{j=1}^{n} \left(T^{-1}\right)_{kj} T_{ji} \right) \mathbf{e}_k$$

De donde

$$\sum_{j=1}^{n} \left(T^{-1}\right)_{kj} T_{ji} = \delta_{ki} \quad \text{o} \quad T^{-1} T = I \tag{5.30}$$

es decir, $T^{-1}$ es la inversa de $T$. Por lo tanto, las matrices $T^{-1}$ y $T$ son no singulares.

Si tenemos dos transformaciones de coordenadas

$$x \xrightarrow{T_1} x' \quad \text{y} \quad x' \xrightarrow{T_2} \tilde{x}$$

$$x' = T_1 x \quad \text{y} \quad \tilde{x} = T_2 x'$$

entonces la transformada $T$

$$x \xrightarrow{T} \tilde{x}$$

se obtiene como el producto de las transformadas anteriores, $T = T_2 T_1$:

$$T x = \tilde{x} = T_2 x' = T_2 T_1 x \tag{5.31}$$

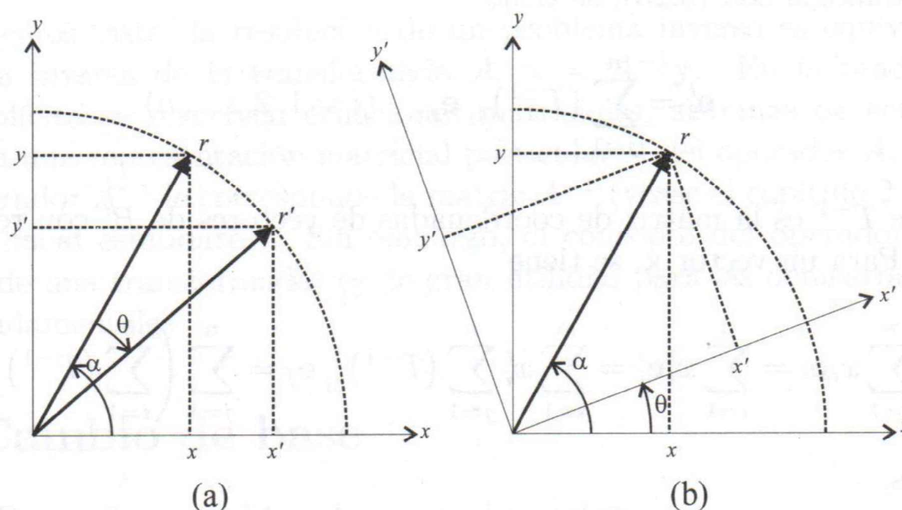



Figura 5.2. Rotación de un vector (a) y de una base (b) en un plano

Nótese que cada transformación lineal $\hat{A}\mathbf{x} = \mathbf{y}$ de $X$ en $X$ puede interpretarse de dos maneras: como una **transformación activa** que transforma los vectores del espacio $X$ con respecto a una base $B = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$, permaneciendo ésta sin cambios; o como una **transformación pasiva** que cambia una base $B$ sin transformar los vectores del espacio $X$, ésta es la transformación de coordenadas de vectores.

**Ejemplo 5.11.** *Transformación de rotación en un plano. En la interpretación activa, se supone que está dada una base cartesiana y el plano se rota a un ángulo $-\theta$ con respecto al centro del sistema de coordenadas. Para determinar la transformación de coordenadas correspondiente a esta operación, se considera la rotación de un vector $\mathbf{r} = (x, y)$ a un ángulo $-\theta$ (rotación al ángulo $\theta$ en el sentido de las manecillas del reloj) con respecto al centro de un sistema cartesiano de coordenadas (véase la figura 5.2(a)). Sean $r = (x, y)$ y $r' = (x', y')$, donde $(x, y)$ y $(x', y')$ son coordenadas de este vector antes y después de su rotación, y sea $\alpha$ el ángulo que forma el vector $\mathbf{r}$ con el eje $x$. Entonces, $|\mathbf{r}| = |\mathbf{r}'| = r$ y $x = r\cos\alpha$, $y = r\,\mathrm{sen}\,\alpha$. De la figura 5.2(a), las coordenadas de dicho vector después de la rotación son*

$$\begin{aligned} x' &= r\cos(\alpha-\theta) = x\cos(-\theta) - y\,\mathrm{sen}(-\theta) \\ y' &= r\,\mathrm{sen}(\alpha-\theta) = x\,\mathrm{sen}(-\theta) + y\cos(-\theta) \end{aligned}$$

*En la forma matricial,*

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos(-\theta) & -\mathrm{sen}(-\theta) \\ \mathrm{sen}(-\theta) & \cos(-\theta) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{5.32}$$

*de donde la matriz de transformación de coordenadas es*

$$R = \begin{pmatrix} \cos\theta & \mathrm{sen}\,\theta \\ -\mathrm{sen}\,\theta & \cos\theta \end{pmatrix}$$

*Nótese que* $\det R = 1$. *Por otra parte, la interpretación pasiva implica la rotación de una base cartesiana a un ángulo* $\theta$ *con respecto al centro de éste, estando el plano fijo. Comparando las figuras 5.2(a) y (b), es fácil ver que la rotación del sistema de coordenadas un ángulo* $\theta$ *en el sentido de las manecillas del reloj es completamente equivalente a la rotación del vector un ángulo* $-\theta$. *De esta manera se demuestra que las coordenadas del vector en el nuevo sistema de coordenadas se expresan mediante la misma fórmula, ecuación (5.32).*

## 5.3.2. Cambio de representación del operador lineal

Considérense las dos bases $B = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ y $B' = \{\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n\}$ en un espacio lineal $n$-dimensional $X$. Sea $\hat{A}$ un operador lineal que mapea $X$ en $X$, $\mathbf{y} = \hat{A}\mathbf{x}$.

Sean

$$\mathbf{y} = \sum_{i=1}^{n} y_i \mathbf{e}_i, \quad \mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i$$

y sea $A = [a_{ij}]$ $(i, j = 1, 2, \ldots, n)$ la representación matricial de $\hat{A}$ en la base $B$,

$$\vec{y} = A\vec{x}$$

Si cambiamos de $B$ a $B'$,

$$\mathbf{y} = \sum_{i=1}^{n} y'_i \mathbf{e}'_i, \quad \mathbf{x} = \sum_{i=1}^{n} x'_i \mathbf{e}'_i$$

Las coordenadas de vectores $\mathbf{y}$ y $\mathbf{x}$ en las bases $B$ y $B'$ están relacionadas entre sí mediante la ecuación (5.25),

$$\vec{x}' = T\vec{x} \quad \text{y} \quad \vec{y}' = T\vec{y}$$

de donde,

$$\vec{y}' = T\vec{y} = TA\vec{x} = TAT^{-1}\vec{x}'$$

Puesto que

$$\vec{y}' = A'\vec{x}'$$

en donde $A'$ es la representación matricial de $\hat{A}$ en la base $B'$, tenemos

$$A' = TAT^{-1} \tag{5.33}$$

Entonces, hemos demostrado que dos matrices $A'$ y $A$, que son representaciones del mismo operador $\hat{A}$ en dos diferentes bases $B'$ y $B$, están relacionadas entre sí mediante una transformación de semejanza, siendo $T$ la matriz de transformación de coordenadas al cambiar la base $B$ a $B'$. En otras palabras, a cada operador lineal $\hat{A}$ en $X$ le corresponde una clase de matrices semejantes; estas matrices representan el operador dado $\hat{A}$ en diferentes bases.

## 5.4. Métrica y ortogonalidad

El concepto de la longitud de un vector y del ángulo entre dos vectores del espacio físico tridimensional está contenido en el concepto del producto escalar. Este último caracteriza plenamente los vectores del espacio físico. En cuanto a los espacios vectoriales arbitrarios, que tienen la estructura matemática más compleja que el espacio físico, es de gran utilidad la generalización del concepto de producto escalar. Un producto escalar en un espacio lineal arbitrario determina la longitud de cada vector y el ángulo entre dos vectores, es decir, la métrica de este espacio. En un espacio con métrica se puede introducir el concepto de ortogonalidad entre dos vectores y las bases ortonormales. Los operadores lineales, considerados en las bases ortonormales, conducen de manera natural al estudio de propiedades de clases especiales de matrices (normales, hermitianas, unitarias, etcétera).

### 5.4.1. Producto interior

En esta sección hablamos de una generalización del concepto de producto escalar.

**Definición 5.9.** *Sea $X$ un espacio vectorial complejo. Un* ***producto interno (interior, escalar)*** *en $X$ es una función que asigna a cada par ordenado de vectores $\mathbf{x}$ y $\mathbf{y}$ un número complejo que se denota por $(\mathbf{x},\mathbf{y})$ y satisface las siguientes condiciones, para toda $\mathbf{x}$, $\mathbf{y}$ y $\mathbf{z}$ en $X$ y cualquier número $\alpha$ complejo:*

1. $(\mathbf{x},\mathbf{y}) = (\mathbf{y},\mathbf{x})^*$,
2. $(\alpha\mathbf{x},\mathbf{y}) = \alpha(\mathbf{x},\mathbf{y})$,
3. $(\mathbf{x}+\mathbf{y},\mathbf{z}) = (\mathbf{x},\mathbf{z}) + (\mathbf{y},\mathbf{z})$,
4. $(\mathbf{x},\mathbf{x}) \geq 0$, $\mathbf{x} = 0$ si y sólo si $(\mathbf{x},\mathbf{x}) = 0$.

En tal caso, se dice que el espacio $X$ tiene una **métrica hermitiana positiva** y

$$|\mathbf{x}| = \sqrt{(\mathbf{x},\mathbf{x})} \tag{5.34}$$

es una **norma** (longitud) de $\mathbf{x}$ inducida por el producto interior. El espacio $X$ con una métrica hermitiana positiva recibe el nombre de **espacio vectorial unitario**. Se dice que un vector $\mathbf{n}$ es normalizado si su longitud $|\mathbf{n}| = 1$. Es claro que para cualquier vector $\mathbf{x}$, el vector $\mathbf{n} = \mathbf{x}/|\mathbf{x}|$ es normalizado.

Nótese que de la definición de producto escalar, incisos 1, 2 y 3, se tiene que

$$(\mathbf{x},\alpha\mathbf{y}) = (\alpha\mathbf{y},\mathbf{x})^* = \alpha^*(\mathbf{x},\mathbf{y}) \tag{5.35}$$

y también

$$(\mathbf{x},\mathbf{y}+\mathbf{z}) = (\mathbf{y}+\mathbf{z},\mathbf{x})^* = (\mathbf{x},\mathbf{y}) + (\mathbf{x},\mathbf{z}) \tag{5.36}$$

En un espacio real

$$(\mathbf{x},\mathbf{y}) = (\mathbf{y},\mathbf{x})$$
$$(\alpha\mathbf{x},\mathbf{y}) = \alpha(\mathbf{x},\mathbf{y}) = (\mathbf{x},\alpha\mathbf{y})$$

en donde $\alpha$ ya es cualquier número real.

**Ejemplo 5.12.** *Sea $B$ una base en un espacio $n$-dimensional $X$ y sean $\mathbf{x}$ y $\mathbf{y}$ dos vectores cualesquiera en $X$ con coordenadas $(x_1, x_2, \ldots, x_n)$ y $(y_1, y_2, \ldots, y_n)$, respectivamente. Entonces, se conoce como producto interno estándar en $X$ al producto*

$$(\mathbf{x},\mathbf{y}) = x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

*que es similar al producto escalar de dos vectores en el espacio físico en una base ortonormal.*

**Ejemplo 5.13.** *Sea $P^{(n)}$ el espacio vectorial real de polinomios de grado menor que $n$. Se puede comprobar que*

$$(\mathbf{x},\mathbf{y}) = \int_0^1 \mathbf{x}(t)\,\mathbf{y}(t)\,dt$$

*es un producto interior, en donde $\mathbf{x}(t)$ y $\mathbf{y}(t)$ son dos polinomios en $P^{(n)}$.*

Con base en los axiomas de producto interior y la definición de norma, demostraremos a continuación la desigualdad de Schwartz

$$|(\mathbf{x},\mathbf{y})| \leq |\mathbf{x}|\,|\mathbf{y}| \tag{5.37}$$

y, a partir de ésta, la desigualdad del triángulo

$$|\mathbf{x}+\mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}| \tag{5.38}$$

La demostración es como sigue.

Notamos que para toda $\alpha$

$$0 \leq |\mathbf{x}+\alpha\mathbf{y}|^2 = (\mathbf{x},\mathbf{x}) + \alpha(\mathbf{x},\mathbf{y})^* + \alpha^*(\mathbf{x},\mathbf{y}) + \alpha\alpha^*(\mathbf{y},\mathbf{y})$$

Como la desigualdad es verdaderamente válida si $(\mathbf{x},\mathbf{y}) = 0$, suponemos que $(\mathbf{x},\mathbf{y}) \neq 0$ y escogemos $\alpha = -(\mathbf{x},\mathbf{x})/(\mathbf{x},\mathbf{y})$. Combinando términos se obtiene

$$\begin{aligned}
0 &\leq |\mathbf{x}+\alpha\mathbf{y}|^2 \\
&= (\mathbf{x},\mathbf{x}) - (\mathbf{x},\mathbf{x})(\mathbf{x},\mathbf{y})^*/(\mathbf{x},\mathbf{y}) - (\mathbf{x},\mathbf{x})(\mathbf{x},\mathbf{y})/(\mathbf{x},\mathbf{y})^* + \\
&\quad (\mathbf{x},\mathbf{x})^2(\mathbf{y},\mathbf{y})/(\mathbf{x},\mathbf{y})(\mathbf{x},\mathbf{y})^* \\
&= -(\mathbf{x},\mathbf{x}) + ((\mathbf{x},\mathbf{y}) - (\mathbf{x},\mathbf{y})^*)^2/(\mathbf{x},\mathbf{y})(\mathbf{x},\mathbf{y})^* + \\
&\quad (\mathbf{x},\mathbf{x})^2(\mathbf{y},\mathbf{y})/(\mathbf{x},\mathbf{y})(\mathbf{x},\mathbf{y})^* \\
&\leq -(\mathbf{x},\mathbf{x}) + (\mathbf{x},\mathbf{x})^2(\mathbf{y},\mathbf{y})/(\mathbf{x},\mathbf{y})(\mathbf{x},\mathbf{y})^*
\end{aligned}$$

de donde resulta la desigualdad (5.37).

Mediante un cálculo directo, usando (5.37), se tiene

$$\begin{aligned} |\mathbf{x}+\mathbf{y}| &= (\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y})^{1/2} = [(\mathbf{x},\mathbf{x})+2\,\mathrm{Re}\,(\mathbf{x},\mathbf{y})+(\mathbf{y},\mathbf{y})]^{1/2} \\ &\leq [(\mathbf{x},\mathbf{x})+2\,|(\mathbf{x},\mathbf{y})|+(\mathbf{y},\mathbf{y})]^{1/2} \\ &\leq \left[(\mathbf{x},\mathbf{x})+2\,(\mathbf{x},\mathbf{x})^{1/2}\,(\mathbf{y},\mathbf{y})^{1/2}+(\mathbf{y},\mathbf{y})\right]^{1/2} = |\mathbf{x}|+|\mathbf{y}| \end{aligned}$$

que demuestra (5.38).

### 5.4.2. Ortogonalidad y bases ortonormales

Sea $X$ un espacio vectorial $n$-dimensional con un producto escalar y $B = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ una base en $X$. Como

$$\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i, \quad \mathbf{y} = \sum_{i=1}^{n} y_i \mathbf{e}_i$$

en donde $x_i$ y $y_i$ $(i = 1, 2, \ldots, n)$ son coordenadas de $\mathbf{x}$ y $\mathbf{y}$ en la base $B$, se tiene

$$(\mathbf{x}, \mathbf{y}) = \sum_{i,j=1}^{n} h_{ij} x_i y_j^* \tag{5.39}$$

en donde

$$h_{ij} = (\mathbf{e}_i, \mathbf{e}_j) \quad (i, j = 1, 2, \ldots, n) \tag{5.40}$$

En particular,

$$(\mathbf{x}, \mathbf{x}) = \sum_{i,j=1}^{n} h_{ij} x_i x_j^* \geq 0 \tag{5.41}$$

en donde la forma bilineal es no negativa por la definición de producto escalar. También, de la definición de producto interno, se tiene

$$h_{ij} = h_{ji}^* \quad (i, j = 1, 2, \ldots, n) \tag{5.42}$$

es decir, la matriz $H = [h_{ij}]$ es hermitiana, $H^+ = H$. Debido a esto, la forma cuadrática (5.41) se llama hermitiana y la métrica correspondiente es también hermitiana.

Por analogía con el espacio físico tridimensional, se dice que $\mathbf{x}$ y $\mathbf{y}$ son dos **vectores ortogonales** si $(\mathbf{x}, \mathbf{y}) = 0$. Se dice que un sistema de vectores $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m\}$ es ortonormal si $(\mathbf{e}_i, \mathbf{e}_j) = \delta_{ij}$ $(i, j = 1, 2, \ldots, m)$.

Sea $B$ una **base ortonormal**, es decir, $(\mathbf{e}_i, \mathbf{e}_j) = \delta_{ij}$ $(i, j = 1, 2, \ldots, n)$. Sean $x_i$ $(i = 1, 2, \ldots, n)$ coordenadas de $\mathbf{x}$ en la base $B$

$$\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i$$

Entonces, multiplicando la igualdad por $\mathbf{e}_k$ $(k=1,2,\ldots,n)$ por la derecha, se obtiene

$$x_k = (\mathbf{x},\mathbf{e}_k) \quad (k=1,2,\ldots,n)$$
$$\mathbf{x} = \sum_{k=1}^{n} (\mathbf{x},\mathbf{e}_k)\,\mathbf{e}_k \tag{5.43}$$

es decir, en una base ortonormal la coordenada de un vector es igual al producto escalar de este vector por el vector correspondiente de la base. Si $x_i$ y $y_i$ $(i=1,2,\ldots,n)$ son coordenadas de $\mathbf{x}$ y $\mathbf{y}$ en la base ortonormal $B$, entonces, de las ecuaciones (5.39) y (5.41), se tiene

$$(\mathbf{x},\mathbf{y}) = \sum_{i,j=1}^{n} \delta_{ij} x_i y_j^* = \sum_{i=1}^{n} x_i y_i^* \tag{5.44}$$

$$(\mathbf{x},\mathbf{x}) = \sum_{i=1}^{n} x_i x_i^* \tag{5.45}$$

Si $T$ es la matriz de transformación de coordenadas inducida por el cambio de $B=\{\mathbf{e}_i\}$ por $B'=\{\mathbf{e}'_i\}$, suponiendo que $B$ y $B'$ son bases ortonormales, es decir,

$$\mathbf{e}_i = \sum_{j=1}^{n} T_{ji}\mathbf{e}'_j \quad \text{y} \quad \vec{x}' = T\vec{x}$$

entonces

$$\delta_{ik} = (\mathbf{e}_i,\mathbf{e}_k) = \left(\sum_{j=1}^{n} T_{ji}\mathbf{e}'_j, \sum_{m=1}^{n} T_{mk}\mathbf{e}'_m\right) = \sum_{m,j=1}^{n} T_{ji}T_{mk}^* \left(\mathbf{e}'_j,\mathbf{e}'_m\right)$$
$$= \sum_{m,j} T_{ji}T_{mk}^*\delta_{jm} = \sum_j T_{ji}T_{jk}^* = \sum_j \left(T^+\right)_{kj} T_{ji}$$

Finalmente, en forma matricial

$$T^+T = I \quad \Rightarrow \quad T^+ = T^{-1} \tag{5.46}$$

Así, la matriz de transformación de una base ortonormal a otra, también ortonormal, es unitaria. En el caso de un espacio real, la matriz $T$ es real y, por tanto, ortogonal, $T^T = T^{-1}$.

**Problema.** *Sean $B$ y $B'$ dos bases ortonormales. Demuestre que*

$$(\mathbf{x},\mathbf{y}) = \sum_{i=1}^{n} x'_i y_i^{*\prime} = \sum_{k=1}^{n} x_k y_k^*$$

*y, en particular,*

$$(\mathbf{x},\mathbf{x}) = \sum_{i=1}^{n} x'_i x_i^{*\prime} = \sum_{k=1}^{n} x_k x_k^*$$

*en donde las coordenadas de $\mathbf{x}$ y $\mathbf{y}$ en la base $B$ son $x_i$ y $y_i$, en la base $B'$ son $x'_i$ y $y'_i$ $(i=1,2,\ldots,n)$.*

### 5.4.3. Operadores adjuntos, hermitianos y unitarios

Es claro (esto es un postulado) que un cambio de base no debe cambiar las normas de vectores y la orientación mutua de vectores, es decir, que un producto escalar debe ser invariante con respecto a cualquier cambio de la base. Una transformación lineal $\hat{A}$, por lo general, cambia la estructura (métrica) del espacio y el producto escalar de los vectores imágenes. Sin embargo, existe una clase de transformadas lineales (unitarias) que no cambian el producto escalar.

**Definición 5.10.** *Sea $\hat{A}$ una transformación lineal de $X$ a $Y$. Si $X$ tiene un producto interior $(\cdot,\cdot)_X$ y $Y$ tiene un producto interior $(\cdot,\cdot)_Y$ y si existe una transformación lineal $\hat{A}^+$ de $Y$ a $X$ tal que*

$$\left(\hat{A}\mathbf{x},\mathbf{y}\right)_Y = \left(\mathbf{x},\hat{A}^+\mathbf{y}\right)_X \tag{5.47}$$

*para toda $\mathbf{x}\in X$ y $\mathbf{y}\in Y$, entonces $\hat{A}^+$ se llama la **transformación adjunta (operador adjunto)** de $\hat{A}$.*

Un caso particular, pero más frecuente en la práctica, son los operadores adjuntos de las transformaciones lineales de un espacio vectorial $X$ a $X$. Sea $\hat{A}$ un operador lineal de $X$ a $X$. Su operador adjunto $\hat{A}^+$ se define mediante la relación

$$\left(\hat{A}\mathbf{x},\mathbf{y}\right) = \left(\mathbf{x},\hat{A}^+\mathbf{y}\right) \tag{5.48}$$

en donde $\mathbf{x},\mathbf{y}\in X$. Demostramos que para cada tal operador $\hat{A}$ existe sólo un operador adjunto $\hat{A}^+$. Sea $B=\{\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n\}$ una base ortonormal en $X$. Para un vector $\mathbf{y}\in X$ y un supuesto operador adjunto $\hat{A}^+$, a partir de (5.43), se tiene

$$\hat{A}^+\mathbf{y} = \sum_{k=1}^{n}\left(\hat{A}^+\mathbf{y},\mathbf{e}_k\right)\mathbf{e}_k$$

Por la definición (5.48), esta ecuación puede escribirse como

$$\hat{A}^+\mathbf{y} = \sum_{k=1}^{n}\left(\mathbf{y},\hat{A}\mathbf{e}_k\right)\mathbf{e}_k \tag{5.49}$$

Ahora, al tomar la última igualdad en calidad de definición del operador $\hat{A}^+$, es fácil demostrar que $\hat{A}^+$ es lineal y satisface la igualdad (5.48) para toda $\mathbf{x},\mathbf{y}\in X$. Además, la ecuación (5.49) define $\hat{A}^+$ unívocamente. De esta manera comprobamos tanto la existencia como la unicidad del operador adjunto $\hat{A}^+$.

Sea $\hat{A}$ un operador lineal de $X$ a $X$ y $A=[a_{ij}]$ la matriz que corresponde a $\hat{A}$ en una base ortonormal $B=\{\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n\}$. Aplicando la fórmula (5.43) al vector $\hat{A}\mathbf{e}_k=\sum_{i=1}^{n}a_{ik}\mathbf{e}_i$, se tiene

$$a_{ik} = \left(\hat{A}\mathbf{e}_k,\mathbf{e}_i\right) \quad (i,k=1,2,\ldots,n) \tag{5.50}$$

Si al operador $\hat{A}^+$ le corresponde la matriz $A^+ = [a_{ik}^+]$, entonces, de la fórmula (5.50),

$$a_{ik}^+ = \left(\hat{A}^+\mathbf{e}_k, \mathbf{e}_i\right) \quad (i, k = 1, 2, \ldots, n) \tag{5.51}$$

Usando la definición (5.48), de (5.50) y (5.51) se obtiene

$$a_{ik}^+ = a_{ki}^* \quad (i, k = 1, 2, \ldots, n) \tag{5.52}$$

es decir, la matriz $A^+$ es conjugada hermitiana de $A$. Por tanto, en cada base ortonormal a los operadores adjuntos les corresponden matrices conjugadas hermitianas.

De la definición (5.48) es fácil obtener las siguientes propiedades:

$$\left(\hat{A}^+\right)^+ = \hat{A} \tag{5.53}$$

$$\left(\hat{A} + \hat{B}\right)^+ = \hat{A}^+ + \hat{B}^+ \tag{5.54}$$

$$\left(\alpha\hat{A}\right)^+ = \alpha^*\hat{A}^+ \tag{5.55}$$

$$\left(\hat{A}\hat{B}\right)^+ = \hat{B}^+\hat{A}^+ \tag{5.56}$$

que están en plena concordancia con las operaciones de conjugación hermitiana sobre las matrices.

**Definición 5.11.** *Se dice que $\hat{A}$ es un* **operador** *(una* **transformación) lineal normal** *si es conmutativo con el operador adjunto*

$$\hat{A}\hat{A}^+ = \hat{A}^+\hat{A} \tag{5.57}$$

**Definición 5.12.** *Se dice que $\hat{A}$ es un* **operador** *(una* **transformación) hermitiano** *si es igual a su adjunto*

$$\hat{A}^+ = \hat{A} \tag{5.58}$$

**Definición 5.13.** *Se dice que $\hat{A}$ es un* **operador** *(una* **transformación) unitario** *si es inverso a su operador adjunto*

$$\hat{A}\hat{A}^+ = \hat{A}^+\hat{A} = \hat{I} \tag{5.59}$$

Nótese que el operador unitario se puede definir como un **operador isométrico** en el espacio hermitiano (espacio con una métrica positiva), es decir, como un operador que conserva la métrica. En verdad, sea

$$\left(\hat{A}\mathbf{x}, \hat{A}\mathbf{x}'\right) = (\mathbf{x}, \mathbf{x}') \tag{5.60}$$

para dos vectores cualesquiera $\mathbf{x}, \mathbf{x}' \in X$. Entonces, según (5.48), $\left(\mathbf{x}, \hat{A}^+\hat{A}\mathbf{x}'\right) = (\mathbf{x}, \mathbf{x}')$ y, por lo tanto, $\hat{A}^+\hat{A} = \hat{I}$, dado que $\mathbf{x}'$ es un vector arbitrario. En consecuencia, $\hat{A}^+ = \hat{A}^{-1}$. Y viceversa, de (5.59) se obtiene (5.60).

De (5.59) o (5.60) se obtiene que el producto de dos operadores unitarios es un operador unitario también, el operador de identidad $\hat{I}$ es unitario y el operador inverso de un unitario es también un operador unitario. Por tanto, el conjunto de todos los operadores unitarios es un grupo que recibe el nombre de grupo unitario.

Nótese que tanto el operador hermitiano como un operador unitario son a la vez operadores normales.

**Problema.** *Demuestre que en toda base ortonormal la matriz de representación $A$ de un operador hermitiano $\hat{A}$ es hermitiana, $A^+ = A$.*

**Problema.** *Demuestre que en toda base ortonormal la matriz de representación $A$ de un operador unitario $\hat{A}$ es unitaria, $AA^+ = A^+A = I$.*

### 5.4.4. Proyecciones ortogonales y proceso de Gram-Schmidt

Los conjuntos ortogonales tienen una gran importancia tanto en el desarrollo teórico de la Física e Ingeniería como en las aplicaciones computacionales. Debido a ello, en esta sección se presentan algunos procedimientos importantes de proyecciones ortogonales y de ortogonalización. Además, mediante el proceso de ortogonalización de Gram-Schmidt, se demostrará que en cada espacio vectorial $n$-dimensional con una métrica hermitiana existe cuando menos una base ortonormal.

Sea $X$ un espacio vectorial con un producto interior y sea $S$ un subespacio $m$-dimensional de $X$ con una base ortonormal $B = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m\}$, en otras palabras, $B$ puede considerarse como un conjunto generador de $S$. Definimos el **operador de proyección ortogonal** $\hat{P}$ sobre un subespacio $S$ como sigue: para cualquier $\mathbf{x} \in X$

$$\hat{P}\mathbf{x} = \sum_{i=1}^{m} (\mathbf{x}, \mathbf{e}_i)\, \mathbf{e}_i \tag{5.61}$$

El vector $\hat{P}\mathbf{x}$ es la **proyección ortogonal** de $\mathbf{x}$ sobre el subespacio $S$.

Es fácil ver que el operador de proyección es lineal,

1. $\hat{P}(\mathbf{x}_1 + \mathbf{x}_2) = \hat{P}\mathbf{x}_1 + \hat{P}\mathbf{x}_2$ para todos $\mathbf{x}_1, \mathbf{x}_2 \in X$;

2. $\hat{P}(\alpha\mathbf{x}) = \alpha\hat{P}\mathbf{x}$ para todo escalar $\alpha$ y todo $\mathbf{x} \in X$.

Además,

$$\left(\mathbf{x} - \hat{P}\mathbf{x}, \mathbf{x}_S\right) = 0 \tag{5.62}$$

es decir, $\mathbf{x} - \hat{P}\mathbf{x}$ es ortogonal a todo vector $\mathbf{x}_S \in S$. Para demostrarlo, observe que $\mathbf{x} - \hat{P}\mathbf{x}$ es ortogonal a cada vector $\mathbf{e}_k$ $(k = 1, 2, \ldots, m)$:

$$\left(\mathbf{x} - \hat{P}\mathbf{x}, \mathbf{e}_k\right) = (\mathbf{x}, \mathbf{e}_k) - \sum_{i=1}^{m} (\mathbf{x}, \mathbf{e}_i)\,(\mathbf{e}_i, \mathbf{e}_k) = 0$$

Luego, como cada $\mathbf{x}_S \in S$ es una combinación lineal de $\mathbf{e}_k$ $(k=1,2,\ldots,m)$, se obtiene (5.62).

Para cada $\mathbf{x} \in X$, de (5.62) se obtiene la representación unívoca

$$\mathbf{x} = \hat{P}\mathbf{x} + \left(\mathbf{x} - \hat{P}\mathbf{x}\right) \tag{5.63}$$

en donde $\hat{P}\mathbf{x} \in S$ es la proyección ortogonal de $\mathbf{x}$ al subespacio $S$ y $\left(\mathbf{x} - \hat{P}\mathbf{x}\right) \perp S$. La notación $\mathbf{x}_N \perp S$ significa que el vector $\mathbf{x}_N$ es ortogonal al subespacio $S$, es decir, a cada vector del subespacio $S$. Dicha representación es de gran utilidad, dado que al determinar parámetros de modelos matemáticos a menudo es necesario aproximar un vector como una combinación lineal de otros vectores dados. Si los vectores con los cuales se forma la combinación lineal son mutuamente ortogonales, la solución del problema de mejor aproximación se da por el siguiente teorema.

**Teorema 5.2.** *Sea $X$ un espacio vectorial con un producto interior y con una norma inducida por este producto. Sea $S$ un subespacio de $X$ con una base ortonormal de vectores $\mathbf{e}_k$ $(k = 1, 2, \ldots, m)$. Entonces, para cualquier $\mathbf{x} \in X$ y cualquier $\mathbf{x}_S \in S$,*

$$\left|\mathbf{x} - \hat{P}\mathbf{x}\right| \leq |\mathbf{x} - \mathbf{x}_S| \tag{5.64}$$

*es decir, $\hat{P}\mathbf{x}$ es el único punto más cercano en $S$ a $\mathbf{x}$. La igualdad en (5.64) es posible sólo para $\mathbf{x}_S = \hat{P}\mathbf{x}$.*

**Demostración.** *Para cualquier $\mathbf{x} \in X$ y cualquier $\mathbf{x}_S \in S$,*

$$\begin{aligned} |\mathbf{x} - \mathbf{x}_S|^2 &= (\mathbf{x} - \mathbf{x}_S, \mathbf{x} - \mathbf{x}_S) \\ &= \left(\mathbf{x} - \hat{P}\mathbf{x} + \hat{P}\mathbf{x} - \mathbf{x}_S, \mathbf{x} - \hat{P}\mathbf{x} + \hat{P}\mathbf{x} - \mathbf{x}_S\right) \\ &= \left(\mathbf{x} - \hat{P}\mathbf{x}, \mathbf{x} - \hat{P}\mathbf{x}\right) + \left(\hat{P}\mathbf{x} - \mathbf{x}_S, \hat{P}\mathbf{x} - \mathbf{x}_S\right) \end{aligned}$$

*dado que* $\left(\mathbf{x} - \hat{P}\mathbf{x}, \hat{P}\mathbf{x} - \mathbf{x}_S\right) = 0 = \left(\hat{P}\mathbf{x} - \mathbf{x}_S, \mathbf{x} - \hat{P}\mathbf{x}\right)$. *Por tanto,*

$$|\mathbf{x} - \mathbf{x}_S|^2 = \left|\mathbf{x} - \hat{P}\mathbf{x}\right|^2 + \left|\hat{P}\mathbf{x} - \mathbf{x}_S\right|^2$$

*de donde se obtiene (5.64).* ■

Los problemas de mejor aproximación son fáciles de resolver en cuanto se cuenta con una base ortonormal para un subespacio de aproximación $S$. Pero éste no siempre es el caso en las aplicaciones. Sin embargo, en espacios de dimensión finita con un producto interior, partiendo de un conjunto generador cualquiera, siempre se puede construir una base ortonormal de una manera directa. El método que se usa se llama **proceso de Gram-Schmidt**. El proceso de Gram-Schmidt produce una base ortogonal y detecta si el conjunto generador es linealmente dependiente.

Sea $G = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_q\}$ un conjunto generador de un espacio vectorial $n$-dimensional $X$ que tiene un producto escalar, $\mathbf{u}_k \neq 0$ ($n \leq q$, $k = 1, 2, \ldots, q$). El procedimiento tradicional de Gram-Schmidt es como sigue, considerándolo de manera inductiva.

Para construir una base ortogonal $\tilde{B}$, primero se define

$$\tilde{\mathbf{e}}_1 = \mathbf{u}_1 \tag{5.65}$$

En principio, en calidad de $\tilde{\mathbf{e}}_1$ puede escogerse cualquier vector del conjunto $G$.

Después se define

$$\tilde{\mathbf{e}}_2 = \mathbf{u}_2 - \frac{(\mathbf{u}_2, \tilde{\mathbf{e}}_1)}{(\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_1)} \tilde{\mathbf{e}}_1 \tag{5.66}$$

Si $\mathbf{u}_2$ es linealmente dependiente de $\tilde{\mathbf{e}}_1$, $\mathbf{u}_2 = \alpha \tilde{\mathbf{e}}_1$, entonces, de la fórmula anterior sigue que $\tilde{\mathbf{e}}_2 = \mathbf{0}$ y, de esta manera, el vector $\mathbf{u}_2$ está descartado de la construcción de $\tilde{B}$. Es fácil ver que los vectores $\tilde{\mathbf{e}}_2$ y $\tilde{\mathbf{e}}_1$ son ortogonales, $(\tilde{\mathbf{e}}_2, \tilde{\mathbf{e}}_1) = 0$. Nótese que el término $[(\mathbf{u}_2, \tilde{\mathbf{e}}_1) / (\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_1)] \tilde{\mathbf{e}}_1$ representa nada más que la proyección ortogonal de $\mathbf{u}_2$ sobre $\tilde{\mathbf{e}}_1$.

Luego,

$$\tilde{\mathbf{e}}_3 = \mathbf{u}_3 - \frac{(\mathbf{u}_3, \tilde{\mathbf{e}}_1)}{(\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_1)} \tilde{\mathbf{e}}_1 - \frac{(\mathbf{u}_3, \tilde{\mathbf{e}}_2)}{(\tilde{\mathbf{e}}_2, \tilde{\mathbf{e}}_2)} \tilde{\mathbf{e}}_2 \tag{5.67}$$

Si $\mathbf{u}_3$ es linealmente dependiente de $\tilde{\mathbf{e}}_1$ y $\tilde{\mathbf{e}}_2$, $\mathbf{u}_3 = \alpha_1 \tilde{\mathbf{e}}_1 + \alpha_2 \tilde{\mathbf{e}}_2$, entonces, de (5.67) se sigue que $\tilde{\mathbf{e}}_3 = \mathbf{0}$ y, de esta manera, el vector $\mathbf{u}_3$ está descartado de la construcción de $\tilde{B}$. Es fácil ver que $\tilde{\mathbf{e}}_3$ es ortogonal a $\tilde{\mathbf{e}}_2$ y $\tilde{\mathbf{e}}_1$, $(\tilde{\mathbf{e}}_3, \tilde{\mathbf{e}}_1) = 0$ y $(\tilde{\mathbf{e}}_3, \tilde{\mathbf{e}}_2) = 0$. El vector

$$\hat{P}_2 \mathbf{u} \equiv \frac{(\mathbf{u}, \tilde{\mathbf{e}}_1)}{(\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_1)} \tilde{\mathbf{e}}_1 + \frac{(\mathbf{u}, \tilde{\mathbf{e}}_2)}{(\tilde{\mathbf{e}}_2, \tilde{\mathbf{e}}_2)} \tilde{\mathbf{e}}_2$$

representa la proyección ortogonal de $\mathbf{u}$ sobre el subespacio generado por los vectores $\tilde{\mathbf{e}}_1$ y $\tilde{\mathbf{e}}_2$.

De manera similar,

$$\tilde{\mathbf{e}}_k = \mathbf{u}_k - \sum_{i=1}^{k-1} \alpha_{ik} \mathbf{e}_i, \quad (k = 1, 2, \ldots, q) \tag{5.68}$$

en donde $\alpha_{ik} = (\mathbf{u}_k, \tilde{\mathbf{e}}_i) / (\tilde{\mathbf{e}}_i, \tilde{\mathbf{e}}_i)$ si $\tilde{\mathbf{e}}_i \neq \mathbf{0}$ y $\alpha_{ik} = 0$ si $\tilde{\mathbf{e}}_i = \mathbf{0}$. Nótese que $\mathbf{e}_i = \mathbf{0}$ si y sólo si $\mathbf{u}_i$ pertenece al subespacio generado por el conjunto de vectores $\{\mathbf{e}_1, \ldots, \mathbf{e}_{i-1}\}$, $\mathbf{u}_i = \alpha_1 \tilde{\mathbf{e}}_1 + \cdots + \alpha_{i-1} \tilde{\mathbf{e}}_{i-1}$, es decir, cuando los vectores $\{\mathbf{u}_1, \ldots, \mathbf{u}_i\}$ son linealmente dependientes. El vector $\tilde{\mathbf{e}}_k$ es ortogonal a los $\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \ldots, \tilde{\mathbf{e}}_{k-1}$.

Dado que el espacio vectorial $X$ es $n$-dimensional, entonces el conjunto ortogonal $\tilde{B} = \{\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_q\}$ contiene exactamente $n$ vectores distintos del vector $\mathbf{0}$. Dichos vectores forman la base ortogonal $\tilde{B}' = \{\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n\}$. De $\tilde{B}'$, normalizando sus vectores de la forma $\mathbf{e}_i = \tilde{\mathbf{e}}_i / (\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_1)^{1/2}$ $(i = 1, 2, \ldots, n)$, se obtiene la base ortonormal $B = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$.

Para resumir, repetimos que el proceso de Gram-Schmidt produce un conjunto generador ortogonal (una base ortogonal) a partir de un conjunto generador cualquiera y detecta si el conjunto generador original es linealmente dependiente. En cada paso, ecuación (5.68), el procedimiento de ortogonalización produce el conjunto generador ortogonal requerido para el siguiente paso. Una conclusión inmediata del proceso de Gram-Schmidt es que todo espacio vectorial de dimensión finita tiene una base ortonormal.

El proceso tradicional de Gram-Schmidt es elegante, aunque no es bueno computacionalmente. El proceso de Gram-Schmidt modificado y las matrices de Householder son alternativas más eficaces para el cálculo de proyecciones ortogonales (véanse, por ejemplo, [8] y [16]).

**Problema.** *Demuestre que para cada operador de proyección $\hat{P}$, se tiene $\hat{P}^2 = \hat{P}$.*

## 5.5. Eigenvalores y vectores de transformaciones lineales

### 5.5.1. Problema de eigenvalores

Al estudiar la estructura de un operador lineal $\hat{A}$ en un espacio vectorial $X$ ($\hat{A}$ es una transformación lineal de $X$ a $X$) son de importancia fundamental los vectores $\mathbf{x}$ tales que

$$\hat{A}\mathbf{x} = \lambda\mathbf{x} \quad (\mathbf{x} \neq \mathbf{0}) \tag{5.69}$$

Dichos vectores y los números correspondientes $\lambda$ reciben el nombre de **eigenvectores** (vectores propios) y de **eigenvalores** (valores propios o característicos) del operador $\hat{A}$, respectivamente.

Para encontrar eigenvalores y eigenvectores de un operador $\hat{A}$, se escoge una base cualquiera $B = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ en $X$. Sean $\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i$ el vector buscado y $A = [a_{ij}]$ la matriz correspondiente al operador $\hat{A}$ en $B$. Entonces, de la ecuación (5.69), se obtiene el sistema homogéneo de ecuaciones lineales

$$\sum_{j=1}^{n} (a_{ij} - \lambda\delta_{ij})\, x_j = 0 \quad (i = 1, 2, \ldots, n) \tag{5.70}$$

$$A\vec{x} = \lambda I \vec{x} \tag{5.71}$$

con respecto a las coordenadas del vector buscado. El sistema homogéneo de ecuaciones lineales tiene una solución no trivial si y sólo si su determinante es igual a cero

$$\det(A - \lambda I) = 0 \tag{5.72}$$

Esta última se llama **ecuación característica** o **secular** de la matriz $A$ (del operador $\hat{A}$) y es una ecuación algebraica de orden $n$ con respecto a $\lambda$. Al cambiar la base $B$ por alguna otra $B'$, se cambia la representación del operador $\hat{A}$. Las dos matrices, $A'$ y $A$, que son representaciones del mismo operador $\hat{A}$ en dos

diferentes bases, $B'$ y $B$, están relacionadas entre sí mediante la transformación de semejanza, ecuación (5.33),

$$A' = TAT^{-1}$$

siendo $T$ la matriz de transformación de coordenadas al cambiar la base $B$ a $B'$. Usando la igualdad $\det AB = \det BA$, se tiene que

$$\det(A' - \lambda I) = \det\left(TAT^{-1} - \lambda I\right) = \det\left(T(A - \lambda I)T^{-1}\right) = \det(A - \lambda I)$$

y, en consecuencia, la ecuación característica para la matriz $A'$ es idéntica a la de $A$, resultando que los eigenvalores de las matrices semejantes son los mismos. Entonces, la ecuación característica no depende de la base y, por tanto, el conjunto de eigenvalores es una característica propia del operador $\hat{A}$ que no depende de su representación matricial.

Cada eigenvalor del operador $\hat{A}$ es una raíz de la ecuación característica (5.72). Y, al contrario, si $\lambda$ es una raíz de (5.72), entonces, para este $\lambda$, el sistema de ecuaciones (5.70) tiene una solución $\vec{x} = (x_1, x_2, \ldots, x_n)$, es decir, a cada $\lambda$ le corresponde un eigenvector $\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i$ del operador $\hat{A}$. De donde se sigue que cada operador $\hat{A}$ en $X$ tiene no más que $n$ valores propios distintos, en donde $n$ es la dimensionalidad del espacio.

Sean $\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_m$ eigenvectores de $\hat{A}$ correspondientes al mismo valor característico $\lambda$,

$$\hat{A}\mathbf{x}_i = \lambda \mathbf{x}_i \quad (i = 1, 2, \ldots, m)$$

Entonces, cualquier combinación lineal de los vectores $\mathbf{x}_i$, si no es igual a cero, es un eigenvector de $\hat{A}$ con el mismo $\lambda$. En realidad, sea

$$\mathbf{y} = \sum_{i=1}^{m} \alpha_i \mathbf{x}_i$$

en donde $\alpha_i$ son escalares arbitrarios. Entonces,

$$\hat{A}\mathbf{y} = \sum_{i=1}^{m} \alpha_i \hat{A}\mathbf{x}_i = \lambda \sum_{i=1}^{m} \alpha_i \mathbf{x}_i = \lambda \mathbf{y}$$

Por lo tanto, los eigenvectores linealmente independientes, correspondientes al mismo eigenvalor $\lambda$, forman la base de un subespacio propio; cada vector de este subespacio es un eigenvector de $\hat{A}$ correspondiente al mismo $\lambda$. En particular, cada eigenvector genera un subespacio propio unidimensional. No obstante, una combinación lineal de eigenvectores de $\hat{A}$ correspondientes a diferentes valores propios, en general, no es un eigenvector de $\hat{A}$.

### 5.5.2. Diagonalización

**Lema 5.1** *Los eigenvectores correspondientes a diferentes valores propios son linealmente independientes.*

**Demostración.** *Sea*

$$\hat{A}\mathbf{x}_i = \lambda_i \mathbf{x}_i \quad (\mathbf{x}_i \neq \mathbf{0};\ \lambda_i \neq \lambda_k \ si\, i \neq k;\ i,k = 1,2,\ldots,m)$$

*Suponemos que*

$$\sum_{i=1}^{m} \alpha_i \mathbf{x}_i = \mathbf{0} \tag{5.73}$$

*de donde, aplicando $\hat{A} - \lambda_1 \hat{I}$ a ambas partes de esta igualdad, se tiene*

$$\sum_{i=2}^{m} \alpha_i (\lambda_i - \lambda_1)\, \mathbf{x}_i = \mathbf{0}$$

*Aplicando a ésta los operadores $\hat{A} - \lambda_2 \hat{I}$, $\hat{A} - \lambda_3 \hat{I}$,..., $\hat{A} - \lambda_{m-1} \hat{I}$ consecutivamente, se obtiene*

$$\alpha_m (\lambda_m - \lambda_{m-1})(\lambda_m - \lambda_{m-2}) \cdots (\lambda_m - \lambda_1)\, \mathbf{x}_m = \mathbf{0}$$

*de donde $\alpha_m = 0$. De manera similar, aplicando a (5.73) las secuencias de $(m-1)$ operadores $\hat{A} - \lambda_i \hat{I}$ $(i = 1,2,\ldots,m)$, de manera que cada secuencia tiene un solo operador faltante, se obtiene*

$$\alpha_1 = \alpha_2 = \cdots = \alpha_m = 0$$

*lo que demuestra la independencia lineal de los vectores $\mathbf{x}_i$ $(i = 1,2,\ldots,m)$.* ■

Si la ecuación característica del operador $\hat{A}$ tiene $m\,(\leq n)$ diferentes raíces (eigenvalores), entonces, según el lema, los eigenvectores correspondientes a estos valores propios son linealmente independientes.

Se dice que $\hat{A}$ en $X$ es un **operador de estructura simple**, si $\hat{A}$ tiene $n$ eigenvectores linealmente independientes, en donde $n$ es la dimensionalidad de $X$. Por lo tanto, $\hat{A}$ tiene la estructura simple, si su ecuación característica tiene $n$ raíces distintas. Dicha condición es suficiente, pero no es necesaria. Existen los operadores de estructura simple cuya ecuación característica tiene raíces múltiples. En el caso de un operador de estructura simple, sus eigenvectores forman una base en $X$.

**Lema 5.2** *Cada operador de estructura simple tiene una representación matricial diagonal en la base de sus eigenvectores, siendo los elementos diagonales de esta matriz los valores propios del operador.*

**Demostración 1.** *Sea $\hat{A}$ un operador de estructura simple en un espacio $n$-dimensional $X$ y sea $\{\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_n\}$ un conjunto de $n$ eigenvectores linealmente independientes, correspondientes a los eigenvalores $\lambda_1, \lambda_2, \ldots, \lambda_n$,*

$$\hat{A}\mathbf{x}_i = \lambda_i \mathbf{x}_i \quad (i = 1,2,\ldots,n)$$

*En general, algunos de los eigenvalores pueden ser múltiplos. Los eigenvectores* $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ *forman una base* $B_\lambda$ *en* $X$. *Por la definición de matriz de representación* $A = [a_{ij}]$ *del operador* $\hat{A}$ *en la base* $B_\lambda$, *se tiene*

$$\hat{A}\mathbf{x}_j = \sum_{i=1}^{n} a_{ij}\mathbf{x}_i = \lambda_j \mathbf{x}_j$$

*Entonces, se tienen combinaciones lineales*

$$\sum_{i=1}^{n} (a_{ij} - \lambda_j \delta_{ij})\, \mathbf{x}_i = \mathbf{0} \quad (j = 1, 2, \ldots, n)$$

*Debido a que los eigenvectores* $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ *son linealmente independientes, los coeficientes de cada combinación lineal son ceros. Entonces, se obtiene*

$$a_{ij} = \lambda_j \delta_{ij} \quad (i, j = 1, 2, \ldots, n) \tag{5.74}$$

*con lo que se completa la demostración.* ■

**Demostración 2.** *En términos de matrices, el lema puede demostrarse de la siguiente manera. En una base* $B$, *sea* $A$ *la matriz de representación de* $\hat{A}$ *y sea* $\vec{x}_j = \left(x_1^j, x_2^j, \ldots, x_n^j\right)$ *el eigenvector columna de coordenadas de eigenvector* $\mathbf{x}_j$, *es decir,* $A\vec{x}_j = \lambda_j \vec{x}_j$ $(j = 1, 2, \ldots, n)$. *Formamos la matriz* $P = [\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n]$ *cuyas columnas son eigenvectores columnas* $\vec{x}_j$. *Sea* $\Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$ *la matriz diagonal con los eigenvalores asociados. Entonces, de* $A\vec{x}_j = \lambda_j \vec{x}_j$ *se obtiene que* $AP = [\lambda_1 \vec{x}_1, \lambda_2 \vec{x}_2, \ldots, \lambda_n \vec{x}_n] = P\Lambda$. *Como* $P$ *es no singular, existe su inversa* $P^{-1}$ *y, por tanto,*

$$\Lambda = P^{-1}AP \tag{5.75}$$

*Identificando a* $T = P^{-1}$ *como la matriz de transformación de coordenadas por el cambio de base* $B$ *a* $B_\lambda = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$, *se tiene que* $\Lambda$ *es la representación de* $\hat{A}$ *en la base* $B_\lambda$ *(véanse las ecuaciones (5.23), (5.27) y (5.33)) completando con esto la demostración.* ■

### 5.5.3. Espectro de operadores normales, hermitianos y unitarios

Ante todo mostramos una propiedad importante de los operadores conmutativos.

**Lema 5.3** *Dos operadores conmutativos,* $\hat{A}$ *y* $\hat{B}$ $\left(\hat{A}\hat{B} = \hat{B}\hat{A}\right)$, *cuando menos tienen un eigenvector en común.*

**Demostración.** *Sea* $\mathbf{x}$ *un eigenvector de* $\hat{A}$, $\hat{A}\mathbf{x} = \lambda\mathbf{x}$ *y* $\mathbf{x} \neq \mathbf{0}$. *Entonces, por la conmutatividad de* $\hat{A}$ *y* $\hat{B}$,

$$\hat{A}\hat{B}^k\mathbf{x} = \lambda\hat{B}^k\mathbf{x} \tag{5.76}$$

*Existe un número $m \leq n$ tal que en la serie de los vectores*

$$\mathbf{x},\ \hat{B}\mathbf{x},\ \hat{B}^2\mathbf{x}, \ldots$$

*los primeros $m$ vectores son linealmente independientes, mientras que el vector $\hat{B}^m\mathbf{x}$ ya es una combinación lineal de los anteriores. Entonces, el subespacio $S$ generado por los vectores $\left\{\mathbf{x}, \hat{B}\mathbf{x}, \ldots, \hat{B}^{m-1}\mathbf{x}\right\}$ es invariante con respecto a $\hat{B}$ y, por lo tanto, en $S$ existe un eigenvector $\mathbf{u}$ de $\hat{B}$, $\hat{B}\mathbf{u} = \mu\mathbf{u}$, $\mathbf{u} \neq \mathbf{0}$. Por otro lado, los vectores $\left\{\mathbf{x}, \hat{B}\mathbf{x}, \ldots, \hat{B}^{m-1}\mathbf{x}\right\}$ son eigenvectores de $\hat{A}$ correspondientes al mismo eigenvalor $\lambda$, ecuación (5.76). Entonces, cualquier combinación lineal de estos vectores, en particular el vector $\mathbf{u}$, también es un eigenvector de $\hat{A}$ correspondiente al valor propio $\lambda$. Esto completa la demostración.* ■

Sea $\hat{A}$ un operador normal en un espacio hermitiano $n$-dimensional $X$. En este caso, los operadores $\hat{A}$ y $\hat{A}^+$ son conmutativos y, por tanto, tienen un eigenvector $\mathbf{x}_1$ en común

$$\hat{A}\mathbf{x}_1 = \lambda_1\mathbf{x}_1, \quad \hat{A}^+\mathbf{x}_1 = \lambda_1^*\mathbf{x}_1$$

en donde el eigenvalor $\lambda_1^*$ de $\hat{A}^+$ es complejo conjugado de $\lambda_1$. Sea $S_1 = [\mathbf{x}_1]$ el subespacio unidimensional generado por $\mathbf{x}_1$ y sea $T_1$ su complemento ortogonal en $X$:

$$X = S_1 + T_1, \quad S_1 \perp T_1$$

Dado que $S_1$ es invariante con respecto a $\hat{A}$ y $\hat{A}^+$, el subespacio $T_1$ es también invariante con respecto a éstos. Por tanto, según el lema, los operadores conmutativos $\hat{A}$ y $\hat{A}^+$ tienen en común un eigenvector $\mathbf{x}_2$ en $T_1$,

$$\hat{A}\mathbf{x}_2 = \lambda_2\mathbf{x}_2, \quad \hat{A}^+\mathbf{x}_2 = \lambda_2^*\mathbf{x}_2$$

Por supuesto, $\mathbf{x}_1 \perp \mathbf{x}_2$. Si $S_2 = [\mathbf{x}_1, \mathbf{x}_2]$ es el subespacio bidimensional generado por $\mathbf{x}_1$ y $\mathbf{x}_2$, entonces

$$X = S_2 + T_2, \quad S_2 \perp T_2$$

y, de manera similar, se establece la existencia en $T_2$ de un eigenvector $\mathbf{x}_3$ de $\hat{A}$ y $\hat{A}^+$. Además, se tiene que $\mathbf{x}_1 \perp \mathbf{x}_3$ y $\mathbf{x}_2 \perp \mathbf{x}_3$. Siguiendo el proceso, se obtiene un conjunto de $n$ eigenvectores $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ perpendiculares entre sí de $\hat{A}$ y $\hat{A}^+$:

$$\begin{aligned} \hat{A}\mathbf{x}_i &= \lambda_i\mathbf{x}_i, \quad \hat{A}^+\mathbf{x}_i = \lambda_i^*\mathbf{x}_i \\ (\mathbf{x}_i, \mathbf{x}_k) &= 0 \quad \text{para} \quad i \neq k \quad (i, k = 1, 2, \ldots, n) \end{aligned} \tag{5.77}$$

Los vectores $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ pueden ser normalizados. De esta manera demostramos que todo operador normal en un espacio hermitiano $n$-dimensional tiene un sistema ortonormal de $n$ eigenvectores. Además, de (5.77) se sigue que los operadores $\hat{A}$ y $\hat{A}^+$ tienen los mismos eigenvectores.

Ahora, suponemos que $\hat{A}$ tiene un sistema completo de eigenvectores ortonormales

$$\hat{A}\mathbf{x}_i = \lambda_i\mathbf{x}_i, \quad (\mathbf{x}_i, \mathbf{x}_k) = \delta_{ik} \quad (i, k = 1, 2, \ldots, n)$$

Demostramos que $\hat{A}$ es un operador normal. Definimos

$$\mathbf{y}_i = \hat{A}^+\mathbf{x}_i - \lambda_i^*\mathbf{x}_i$$

Luego,

$$\begin{aligned}(\mathbf{x}_k, \mathbf{y}_i) &= \left(\mathbf{x}_k, \hat{A}^+\mathbf{x}_i\right) - \lambda_i (\mathbf{x}_k, \mathbf{x}_i) = \left(\hat{A}\mathbf{x}_k, \mathbf{x}_i\right) - \lambda_i (\mathbf{x}_k, \mathbf{x}_i) \\ &= (\lambda_k - \lambda_i)\, \delta_{ik} = 0 \quad (i, k = 1, 2, \ldots, n)\end{aligned}$$

En consecuencia,

$$\mathbf{y}_i = \hat{A}^+\mathbf{x}_i - \lambda_i^*\mathbf{x}_i = \mathbf{0} \quad (i = 1, 2, \ldots, n)$$

y es válida (5.77). Por lo tanto,

$$\hat{A}\hat{A}^+\mathbf{x}_i = \lambda_i\lambda_i^*\mathbf{x}_i \quad \text{y} \quad \hat{A}^+\hat{A}\mathbf{x}_i = \lambda_i^*\lambda_i\mathbf{x}_i$$

de donde

$$\hat{A}\hat{A}^+ = \hat{A}^+\hat{A}$$

De esta manera completamos la demostración de un teorema importante.

**Teorema 5.3.** *Un operador lineal es normal si y sólo si tiene un sistema completo de eigenvectores ortonormales.*

En particular, demostramos que en todo caso el operador normal es un operador de estructura simple, es decir, tiene un sistema de $n$ eigenvectores linealmente independientes.

A continuación considérese el espectro de operadores hermitianos y unitarios.

**Teorema 5.4.** *Un operador lineal $\hat{H}$ es hermitiano si y sólo si tiene un sistema completo de eigenvectores ortonormales con eigenvalores reales.*

**Demostración.** *Como el operador hermitiano $\hat{H}$ es un caso particular de los operadores normales, éste tiene un sistema ortonormal de eigenvectores y, de (5.77), se tiene*

$$\hat{H}\mathbf{x}_i = \lambda_i\mathbf{x}_i, \quad \hat{H}^+\mathbf{x}_i = \lambda_i^*\mathbf{x}_i, \quad (\mathbf{x}_i, \mathbf{x}_k) = \delta_{ik} \quad (i, k = 1, 2, \ldots, n)$$

*Luego, restando*

$$(\lambda_i - \lambda_i^*)\,\mathbf{x}_i = \hat{H}\mathbf{x}_i - \hat{H}^+\mathbf{x}_i = \hat{H}\mathbf{x}_i - \hat{H}\mathbf{x}_i = \mathbf{0}$$

*y como $\mathbf{x}_i \neq \mathbf{0}$, entonces*

$$\lambda_i = \lambda_i^* \quad (i = 1, 2, \ldots, n) \tag{5.78}$$

*Los eigenvalores del operador hermitiano son reales.*

*Al contrario, un operador normal $\hat{H}$ con eigenvalores reales en todo caso es hermitiano. En realidad, si se tiene*

$$\hat{H}\mathbf{x}_i = \lambda_i \mathbf{x}_i, \quad \hat{H}^+\mathbf{x}_i = \lambda_i \mathbf{x}_i, \quad (\mathbf{x}_i, \mathbf{x}_k) = \delta_{ik} \quad (i, k = 1, 2, \ldots, n)$$

*entonces*

$$\hat{H}\mathbf{x}_i = \hat{H}^+\mathbf{x}_i \quad (i = 1, 2, \ldots, n)$$

*de donde $\hat{H} = \hat{H}^+$, completando la demostración.* ■

Para enfatizar una propiedad importante de los operadores hermitianos, comprobamos el siguiente teorema.

**Teorema 5.5.** *Si los eigenvalores de un operador hermitiano $\hat{H}$ son diferentes, los eigenvectores correspondientes son ortogonales.*

**Demostración.** *Sean $\mathbf{x}_1$ y $\mathbf{x}_2$ dos eigenvectores de $\hat{H}$ correspondientes a dos diferentes eigenvalores $\lambda_1 \neq \lambda_2$,*

$$\hat{H}\mathbf{x}_1 = \lambda_1 \mathbf{x}_1, \quad \hat{H}\mathbf{x}_2 = \lambda_2 \mathbf{x}_2$$

*De donde*

$$\left(\hat{H}\mathbf{x}_1, \mathbf{x}_2\right) = \lambda_1 (\mathbf{x}_1, \mathbf{x}_2), \quad \left(\mathbf{x}_1, \hat{H}\mathbf{x}_2\right) = \lambda_2 (\mathbf{x}_1, \mathbf{x}_2)$$

*Restando éstas,*

$$\begin{aligned} (\lambda_1 - \lambda_2)(\mathbf{x}_1, \mathbf{x}_2) &= \left(\hat{H}\mathbf{x}_1, \mathbf{x}_2\right) - \left(\mathbf{x}_1, \hat{H}\mathbf{x}_2\right) \\ &= \left(\mathbf{x}_1, \hat{H}^+\mathbf{x}_2\right) - \left(\mathbf{x}_1, \hat{H}\mathbf{x}_2\right) = \mathbf{0} \end{aligned}$$

*Por lo tanto, $(\mathbf{x}_1, \mathbf{x}_2) = 0$, demostrándose así que $\mathbf{x}_1$ y $\mathbf{x}_2$ son ortogonales.* ■

Dado que cada operador unitario $\hat{U}$ es normal, tiene un sistema completo de eigenvectores ortonormales

$$\hat{U}\mathbf{x}_i = \lambda_i \mathbf{x}_i, \quad \hat{U}^+\mathbf{x}_i = \lambda_i^* \mathbf{x}_i, \quad (\mathbf{x}_i, \mathbf{x}_k) = \delta_{ik} \quad (i, k = 1, 2, \ldots, n) \tag{5.79}$$

Aplicando $\hat{U}\hat{U}^+ = \hat{I}$, de (5.79), se tiene

$$\mathbf{x}_i = \hat{U}\hat{U}^+\mathbf{x}_i = \lambda_i^* \hat{U}\mathbf{x}_i = \lambda_i^* \lambda_i \mathbf{x}_i \quad (i = 1, 2, \ldots, n)$$

de donde

$$\lambda_i^* \lambda_i = 1 \quad (i = 1, 2, \ldots, n) \tag{5.80}$$

Y viceversa, de (5.79) y (5.80) se obtiene $\hat{U}\hat{U}^+ = \hat{I}$, demostrándose así el siguiente teorema.

**Teorema 5.6.** *Un operador lineal $\hat{U}$ es unitario si y sólo si tiene un sistema completo de eigenvectores ortonormales con eigenvalores de módulo igual a uno.*

En cada base ortonormal un operador normal (en particular, hermitiano o unitario) es representado por una matriz normal (hermitiana, unitaria), y viceversa. Entonces, como consecuencia de los teoremas anteriores, se obtienen:

**Teorema 5.7.** *Una matriz $A$ es normal si y sólo si es semejante unitariamente a una matriz diagonal*

$$A = T\,diag(\lambda_1, \lambda_2, \ldots, \lambda_n)\,T^{-1} \quad (T^{+} = T^{-1}) \tag{5.81}$$

**Teorema 5.8.** *Una matriz $H$ es hermitiana si y sólo si es semejante unitariamente a una matriz diagonal con números reales sobre la diagonal*

$$H = T\,diag(\lambda_1, \lambda_2, \ldots, \lambda_n)\,T^{-1} \quad \left(T^{+} = T^{-1},\ \lambda_i = \lambda_i^*,\ i = 1, 2, \ldots, n\right) \tag{5.82}$$

**Teorema 5.9.** *Una matriz $U$ es unitaria si y sólo si es semejante unitariamente a una matriz diagonal con los elementos de módulo igual a uno en la diagonal*

$$U = T\,diag(\lambda_1, \lambda_2, ..., \lambda_n)\,T^{-1} \quad \left(T^{+} = T^{-1},\ \lambda_i\lambda_i^* = 1,\ i = 1, 2, ..., n\right) \tag{5.83}$$

**Problema.** *Demuestre que cada subespacio invariante con respecto a un operador $\hat{A}$ contiene cuando menos un eigenvector de $\hat{A}$.*

**Problema.** *Compruebe que los eigenvalores de $\hat{A}$ y $\hat{A}^{+}$, correspondientes al mismo eigenvector $\mathbf{x}$, son complejos conjugados, $\hat{A}\mathbf{x} = \lambda\mathbf{x}$, $\quad \hat{A}^{+}\mathbf{x} = \lambda^*\mathbf{x}$.*

# 5.6. Espacios lineales de dimensión infinita

En esta sección se considera en forma resumida la extensión del concepto de espacios vectoriales de dimensión finita a los de dimensión infinita que se presentan con bastante frecuencia en la física (mecánica cuántica, electromagnetismo) y en el procesamiento de señales e imágenes. La dimensionalidad infinita de un espacio implica un número infinito de los vectores linealmente independientes. Se distinguen dos tipos de espacios de dimensión infinita: los que tienen una base contable y los que tienen una base "continua". La diferencia entre éstos se aclara a continuación.

## 5.6.1. Espacios de dimensión numerable

Sea $X$ un espacio lineal $n$-dimensional y sea $B = \{\mathbf{e}_i : i = 1, 2, \ldots, n\}$ una base en $X$. Si $\mathbf{x}$ es un vector de $X$ y $\hat{A}$ es una transformación lineal de $X$ a $X$, entonces la representación de $\mathbf{x}$ es el vector columna de coordenadas

$$\vec{x} = (x_1, x_2, \ldots, x_n), \quad \mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i$$

y la matriz

$$A = [a_{ij}], \quad \hat{A}\mathbf{e}_j = \sum_{i=1}^{n} a_{ij}\mathbf{e}_i, \quad (i, j = 1, 2, \ldots, n)$$

Si, de manera formal $n \to \infty$, tenemos un **espacio lineal de dimensión infinita numerable** $X_\infty$. El problema más serio que enfrenta uno al tender $n \to \infty$ es que las representaciones matriciales y las operaciones vectoriales en $X_\infty$ ya contienen las sumas y series infinitas. Por ejemplo, el producto interior y la norma asociada podrían definirse como

$$(\mathbf{x}, \mathbf{y}) = \sum_{i,j=1}^{\infty} h_{ij} x_i y_j^*, \quad (\mathbf{x}, \mathbf{x}) = \sum_{i,j=1}^{\infty} h_{ij} x_i x_j^* \geq 0 \tag{5.84}$$

en donde

$$h_{ij} = (\mathbf{e}_i, \mathbf{e}_j) \quad (i, j = 1, 2, \ldots, n)$$

Sin embargo, estas sumas divergen en general. La solución típica del problema consiste en considerar sólo el subconjunto de vectores de norma finita

$$\tilde{X}_\infty = \left\{ \mathbf{x} : |\mathbf{x}|^2 = (\mathbf{x}, \mathbf{x}) < \infty \right\}$$

Comprobamos que $\tilde{X}_\infty$ es un espacio vectorial y $\tilde{X}_\infty \subset X$, es decir, es cerrado con respecto a la multiplicación de vectores por escalares y la suma de vectores. En realidad, si $\mathbf{x} \in \tilde{X}_\infty$, entonces también $\alpha\mathbf{x} \in \tilde{X}_\infty$, puesto que $|\alpha\mathbf{x}| = |\alpha| |\mathbf{x}| < \infty$ para cualquier número complejo $\alpha$. Se puede comprobar que las series infinitas en (5.84) convergen absolutamente, cuando $\mathbf{x}, \mathbf{y} \in \tilde{X}_\infty$ y $(\mathbf{x}, \mathbf{y})$ satisface los axiomas del producto interior. En este caso es válida la desigualdad de triángulo. Por lo tanto, $|\mathbf{x} + \mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}| < \infty$ y $\mathbf{x} + \mathbf{y} \in \tilde{X}_\infty$, si $\mathbf{x}, \mathbf{y} \in \tilde{X}_\infty$. En consecuencia, $\tilde{X}_\infty$ es un espacio vectorial infinito con una base numerable y $(\mathbf{x}, \mathbf{y})$ define un producto interior en $\tilde{X}_\infty$.

Sea $\mathcal{H}_T$ el conjunto de todas las funciones complejas medibles de una variable real $t$ que son periódicas de periodo $T$, es decir, $f(t + T) = f(t)$ para toda $t \in R$ y se satisface la desigualdad

$$\|f\|_T^2 \equiv \int_{t_0}^{t_0+T} |f(t)|^2 \, dt < \infty \tag{5.85}$$

en donde hemos empleado el símbolo $\|\cdots\|$ para la norma de los elementos de $\mathcal{H}_T$, con el fin de no confundirla con el módulo de la función $f(t)$. La integral de la derecha de (5.85), por supuesto, no depende de $t_0$ debido a la periodicidad de funciones. Definimos el producto interior con la fórmula

$$(f, g)_T = \int_{t_0}^{t_0+T} f(t) \, g^*(t) \, dt \tag{5.86}$$

Con esta definición el conjunto $\mathcal{H}_T$ es nada más que el espacio $L^2(t_0, t_0 + T)$. Las funciones

$$e_n(t) = \exp(in\omega t), \quad \omega = 2\pi/T \quad (n = 0, \pm 1, \pm 2, \ldots) \tag{5.87}$$

son periódicas de periodo $T$ con la norma $\|e_n\|_T^2 = T < \infty$ y, por tanto, pertenecen al espacio $\mathcal{H}_T$. Además, son ortogonales

$$(e_n, e_m)_T = T\delta_{nm} \tag{5.88}$$

y, por tanto, linealmente independientes. Considérese la función $f_N \in \mathcal{H}_T$, que es la combinación lineal de los $e_n(t)$ $(n = 0, \pm 1, \pm 2, \ldots, N)$,

$$f_N(t) = \frac{1}{T} \sum_{n=-N}^{N} f_n e_n(t) \tag{5.89}$$

en donde los coeficientes $f_n$ se determinan por la fórmula

$$f_n = (f, e_n)_T \quad (n = 0, \pm 1, \pm 2, \ldots, N) \tag{5.90}$$

A partir de la definición de los $f_n$ y usando la ortogonalidad de los $e_n(t)$, de (5.89) y (5.90) se tiene

$$f_n = (f, e_n)_T = (f_N, e_n)_T \quad (n = 0, \pm 1, \pm 2, \ldots, N) \tag{5.91}$$

Así, la función $f_N(t)$ es la proyección ortogonal de $f(t)$ sobre el subespacio generado por el conjunto ortogonal $B_N = \{e_n(t) : n = 0, \pm 1, \pm 2, \ldots, N\}$, siendo $f_n$ las coordenadas de $f_N(t)$ en la base $B_N$. En la teoría de las series de Fourier se demuestra que la secuencia de funciones $\{f_1, f_2, f_3, \ldots\}$ converge en $\mathcal{H}_T$ a la función $f(t)$ en el límite $N \to \infty$. Entonces, en $\mathcal{H}_T$ se tiene la representación de $f(t)$

$$f(t) = \frac{1}{T} \sum_{n=-\infty}^{\infty} f_n e_n(t), \quad f_n = (f, e_n)_T = \int_{t_0}^{t_0+T} f(t)\, e^{-in\omega t} dt \tag{5.92}$$

en la base $B_\infty = \{e_n(t) : n = 0, \pm 1, \pm 2, \ldots\}$. Además, la ortogonalidad de la base $B_\infty$ implica que

$$\|f\|_T^2 = \frac{1}{T} \sum_{n=-\infty}^{\infty} |f_n|^2 < \infty \tag{5.93}$$

La expansión (5.92) se conoce como la serie de Fourier de $f(t) \in \mathcal{H}_T$. La representación de funciones periódicas en la forma de la serie de Fourier asocia a cada función un conjunto infinito numerable de coeficientes de la serie (coordenadas de la función en la base $B_\infty$). Usando la ortogonalidad de la base, se obtiene el producto interior de la forma

$$\begin{aligned}
(f, g)_T &= \int_{t_0}^{t_0+T} f(t)\, g^*(t)\, dt \\
&= \int_{t_0}^{t_0+T} \left( \frac{1}{T} \sum_{n=-\infty}^{\infty} f_n e_n(t) \right) \left( \frac{1}{T} \sum_{m=-\infty}^{\infty} g_m^* e_m^*(t) \right) dt \\
&= \frac{1}{T^2} \sum_{n,m=-\infty}^{\infty} f_n g_m^* (e_n, e_m) = \frac{1}{T} \sum_{n=-\infty}^{\infty} f_n g_n^*
\end{aligned} \tag{5.94}$$

en donde $g_m$ son los coeficientes de la serie de Fourier de $g(t)$.

Las series de Fourier tienen una interpretación bien conocida. Si $f(t)$ se considera como una señal periódica, entonces la expansión (5.92) representa una

descomposición de $f(t)$ en una combinación lineal de modos armónicos $e_n(t)$ de frecuencias $\omega_n = n\omega = n2\pi/T$ $(n = 0, \pm 1, \pm 2, \ldots)$, siendo éstas múltiplos de la frecuencia fundamental $\omega = 2\pi/T$. Para una $f(t) \in \mathcal{H}_T$ dada, el cálculo de los coeficientes de Fourier $f_n$ se llama el análisis, porque éstos representan el contenido de armónicas en $f(t)$. Para el conjunto $\{f_n : n = 0, \pm 1, \pm 2, \ldots\}$ dado, la representación (5.92) se llama la síntesis, porque ésta reproduce la señal a partir de los modos armónicos. En la ecuación (5.93), $\|f\|_T^2$, interpretada como la energía de la señal por un periodo $T$, es distribuida entre los modos armónicos de tal manera que el modo $n$-ésimo contiene la energía $|f_n|^2/T$ por un periodo $T$. Todos los conceptos de procesamiento de una señal mediante las series de Fourier pueden extenderse a otros tipos de análisis y síntesis, reemplazando los vectores $\{e_n(t) : n = 0, \pm 1, \pm 2, \ldots\}$ por alguna otra base que puede ser formada por uno de los conjuntos de funciones especiales. La selección de una base para la expansión de la señal se determina por el carácter del procesamiento correspondiente.

### 5.6.2. Espacios lineales de funciones

La situación se vuelve aún más complicada cuando el índice de las coordenadas de funciones es continuo. En un conjunto de funciones $\mathcal{H}$, de manera natural se definen el producto interior y la norma asociada a éste:

$$(f,g) \equiv \int_{-\infty}^{\infty} f(t)\, g^*(t)\, dt = (g,f)^*, \quad \|f\|^2 \equiv \int_{-\infty}^{\infty} |f(t)|^2\, dt \tag{5.95}$$

Sin embargo, surgen dos problemas asociados con estas definiciones. Los mencionamos a continuación sin entrar en detalles, dado que muchas de las funciones en las aplicaciones son de bastante buen comportamiento.

Primero, resulta que las integrales en (5.95) no existen para todas las funciones. La razón es que las integrales son impropias y, por tanto, pueden ser divergentes. Además, hay una clase de funciones un tanto específicas para las cuales no existen integrales de Riemann aunque el soporte de las funciones puede ser finito. Dicho problema se resuelve con la teoría de medida e integrales de Lebesgue.

Segundo, si se tiene resuelto el primer problema, puede ocurrir que la fórmula (5.95) aún no defina un verdadero producto interior por la falla del axioma de positividad: $\|f\| > 0$ para toda $f(t) \in \mathcal{H}$ cuando $f(t) \neq 0$, y $\|f\| = 0$ si $f(t) = 0$. Este axioma es fundamental para la definición de la igualdad de dos funciones en un espacio de Hilbert. La solución de dicho problema se da mediante la teoría de medida y la generalización del concepto de igualdad de dos funciones, $f(t) = g(t)$, en cada punto $t$ a la igualdad de éstas casi en cualquier punto, salvo los puntos de medida cero.

Un conjunto de funciones $\mathcal{H}$, resueltos los dos problemas mencionados anteriormente, tiene un producto interior (5.95) bien definido que satisface los axiomas del producto interior como en el caso de los espacios vectoriales de dimensión finita. Esto implica la desigualdad del triángulo y la de Schwartz

$$\|f+g\| \leq \|f\| + \|g\|, \quad |(f,g)| \leq \|f\| \cdot \|g\| \tag{5.96}$$

para toda $f, g \in \mathcal{H}$. La desigualdad del triángulo asegura que $\mathcal{H}$ es cerrado con respecto a la suma de vectores y, por tanto, $\mathcal{H}$ es un espacio vectorial complejo. En este caso $\mathcal{H}$ se llama el espacio de funciones (complejas de variable real) con el cuadrado de módulo integrable y éste se denomina $L^2(\mathbf{R})$.

El espacio $L^2(\mathbf{R})$ es un ejemplo de espacio de Hilbert y tiene gran importancia en las aplicaciones. La definición general de espacio de Hilbert es como sigue. Un **espacio de Hilbert** $\mathcal{H}$ es cualquier espacio vectorial con un producto interior que satisface los axiomas correspondientes y, además, es completo en el sentido de que cualquier secuencia $\{f_1, f_2, f_3, ...\}$ en $\mathcal{H}$, tal que $\|f_n - f_m\| \to 0$ cuando $n, m \to \infty$, converge a una función $f \in \mathcal{H}$, es decir, $\|f - f_n\| \to 0$ cuando $n \to \infty$.

Dados dos espacios de Hilbert, $\mathcal{H}$ y $\mathcal{K}$, y una función (transformación, mapeo), $\hat{A} : \mathcal{H} \to \mathcal{K}$, se dice que $\hat{A}$ es un operador lineal si

$$\hat{A}(\alpha f) = \alpha \hat{A} f \quad \text{y} \quad \hat{A}(f + g) = \hat{A} f + \hat{A} g \tag{5.97}$$

para cualquier número complejo $\alpha$ y toda $f, g \in \mathcal{H}$. Se consideran únicamente los operadores acotados, es decir, existe una constante positiva $C$ tal que $\left\|\hat{A} f\right\|_{\mathcal{K}} < C \|f\|_{\mathcal{H}}$ para toda $f \in \mathcal{H}$. El acotamiento de operadores es necesario únicamente para los espacios de Hilbert de dimensión infinita; se puede demostrar que cada operador definido en un espacio de Hilbert de dimensión finita $\mathcal{H}$ es acotado independientemente de que el espacio $\mathcal{K}$ sea finito o infinito. La composición $\hat{B}\hat{A}$ de dos operadores acotados $\hat{A}$ y $\hat{B}$ es también un operador acotado. Muchas de las aplicaciones de dichos conceptos se encuentran a menudo en la mecánica cuántica.

La teoría formal de los espacios de Hilbert es un tanto complicada y se encuentra fuera del propósito de este libro. Por lo tanto, para finalizar esta sección nos restringimos a considerar como un ejemplo el espacio $L^2(\mathbf{R})$ y la transformada de Fourier en éste.

Para pasar de las series de Fourier a la transformada de Fourier de manera más natural, escogemos $t_0 = -T/2$, logrando con esto la simetría del intervalo de integración, y luego hacemos tender $T \to \infty$. En este límite las funciones ya no son periódicas, pero satisfacen la condición $\int_{-\infty}^{\infty} |f(t)|^2 \, dt < \infty$. En consecuencia, $\mathcal{H}_T \to L^2(\mathbf{R})$. Para hacer notar la dependencia explícita de los coeficientes de Fourier $f_n$ de $T$ y $n$, escribimos (5.92) en la forma

$$f_n = \int_{-T/2}^{T/2} f(t) \, e^{-in\omega t} dt \equiv \tilde{f}_T(\omega_n)$$

en donde $\omega_n = n\omega = n2\pi/T$ y se define la función

$$\tilde{f}_T(\omega) \equiv \int_{-T/2}^{T/2} f(t) \, e^{-i\omega t} dt$$

para cualquier frecuencia $\omega \in \mathbf{R}$. Cuando $T \to \infty$, $\tilde{f}_T(\omega)$ se convierte en la transformada integral de Fourier

$$\tilde{f}(\omega) \equiv \int_{-\infty}^{\infty} f(t) \, e^{-i\omega t} dt \tag{5.98}$$

Volviendo a $T$ finito y notando que $\Delta\omega_n = \omega_{n+1} - \omega_n = 2\pi/T$, podemos escribir la serie (5.92) como una suma de Riemann

$$f(t) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \tilde{f}_T(\omega_n) e^{i\omega_n t} \Delta\omega_n$$

En el límite $T \to \infty$ esta suma se convierte en la transformada inversa de Fourier

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{f}(\omega) e^{i\omega t} d\omega \tag{5.99}$$

La relación (5.93) también puede escribirse como una suma de Riemann

$$\|f\|_T^2 = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \left|\tilde{f}_T(\omega_n)\right|^2 \Delta\omega_n$$

que en el límite $T \to \infty$ se convierte en una relación integral

$$\|f\|^2 \equiv \int_{-\infty}^{\infty} |f(\omega)|^2 dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left|\tilde{f}(\omega)\right|^2 d\omega \tag{5.100}$$

conocida como el teorema de Parseval. La última relación demuestra que $f(t) \in L^2(\mathbf{R})$ si y sólo si $\tilde{f}(\omega) \in L^2(\mathbf{R})$.

Nótese que en la transición de las series a las integrales de Fourier, el espacio $\mathcal{H}_T$ se transformó en $L^2(\mathbf{R})$, la base numerable $B_\infty = \{e_n(t) : n = 0, \pm1, \pm2, \ldots\}$ se transformó en el conjunto continuo de funciones $\{e_\omega(t) = e^{i\omega t}\}$ que podría llamarse la base de $L^2(\mathbf{R})$. Sin embargo, las funciones $e_\omega(t) = e^{i\omega t}$ tienen la norma infinita, $\|e_\omega\|^2 = \infty$, y por lo tanto no pertenecen a $L^2(\mathbf{R})$ y no pueden llamarse la base de este espacio. La resolución de dicha contrariedad requiere cuando menos el conocimiento de las funciones generalizadas o distribuciones, que se encuentra fuera de los temas de este libro.

**Nota bibliográfica:** para ampliar los conocimientos sobre el tema de este capítulo, puede consultar [4], [6], [8], [11], [16], [17] y [18].

# Capítulo 6

# Series e integrales de Fourier

## 6.1. Funciones pares e impares, periódicas y ortogonales

### 6.1.1. Funciones pares e impares

Antes de empezar con series de Fourier recordemos algunos de los conceptos básicos de la teoría de funciones. Se dice que una **función** $\varphi(x)$ es **par** si $\varphi(-x) = \varphi(x)$ para toda $x$.

**Ejemplo 6.1.** *Funciones pares son* $x^2$, $|x|$, $\operatorname{sen}^2 x$, $\cos x$.

Una **función** $\chi(x)$ se llama **impar** si $\chi(-x) = -\chi(x)$ para toda $x$.

**Ejemplo 6.2.** *Funciones impares son* $x$, $x^3$, $\operatorname{sen} x$.

Cualquier función $f(x)$ se puede representar como la suma de una función par y otra impar

$$f(x) = \varphi(x) + \chi(x) \tag{6.1}$$

donde

$$\varphi(x) = \frac{1}{2}\left[f(x) + f(-x)\right] = \varphi(-x) \tag{6.2}$$

$$\chi(x) = \frac{1}{2}\left[f(x) - f(-x)\right] = -\chi(-x) \tag{6.3}$$

Es fácil ver que si $\varphi(x)$ es una función par, entonces

$$\int_{-A}^{A} \varphi(x)\, dx = 2\int_{0}^{A} \varphi(x)\, dx \tag{6.4}$$

donde $A$ es cualquier número real.

De manera similar, se comprueba que si $\chi(x)$ es una función impar, entonces

$$\int_{-A}^{A} \varphi(x)\, dx = 0 \tag{6.5}$$

**Problema.** *¿Qué tipo de función será la resultante del producto: a) de dos funciones par; b) de dos funciones impar; c) de una función par por otra impar?*

### 6.1.2. Funciones periódicas

Se dice que $f(x)$ es una **función periódica** si existe un número $T$, positivo por convención, tal que

$$f(x+T) = f(x) \tag{6.6}$$

para toda $x$ real. Este número $T$ se llama el **periodo** de $f(x)$. Si $T$ es un periodo de $f(x)$, entonces un múltiplo de éste, $kT$, también es periodo de esta función

$$f(x+kT) = f((x+(k-1)T)+T) = f(x+(k-1)T) = \cdots = f(x) \tag{6.7}$$

Por lo general, se refiere al periodo mínimo de una función $f(x)$ cuando se habla del periodo de ésta. Si $f(x)$ y $g(x)$ tienen un periodo común $T$, la combinación lineal de éstas

$$h(x) = Af(x) + Bg(x)$$

donde $A$ y $B$ son constantes, también tiene el periodo $T$.

**Ejemplo 6.3.** *Las funciones $f(x) = \operatorname{sen} x$ y $g(x) = \cos 2x$ son periódicas con periodos correspondientes $T_1 = 2\pi$ y $T_2 = \pi$. Por tanto, la función $h(x) = A \operatorname{sen} x + B \cos 2x$ es periódica con periodo $T = 2\pi$.*

**Ejemplo 6.4.** *La función $f(x) = const$ es periódica y tiene como su periodo cualquier número real $T$.*

Una función $\tilde{f}(x)$ definida en un intervalo $a < x \leq a+T$, la podemos continuar periódicamente fuera de este intervalo de la siguiente manera:

$$f(x) = \tilde{f}(x-kT) \quad \text{si} \quad a+kT < x \leq a+(k+1)T \tag{6.8}$$

obteniendo finalmente la función periódica $f(x)$ con periodo $T$, donde los $k$ son números enteros (véase la figura 6.1).

En el caso de una función $\tilde{f}(x)$ definida en un intervalo $0 < x \leq T/2$, la podemos continuar periódicamente resultando una función par o función impar, dependiendo de la manera de continuación. Para obtener la función periódica par, los pasos son los siguientes (véase la figura 6.2):

$$\tilde{\varphi}(x) = \begin{cases} \tilde{f}(x) & \text{si} \quad 0 < x \leq T/2 \\ \tilde{f}(-x) & \text{si} \quad -T/2 < x \leq 0 \end{cases} \tag{6.9}$$

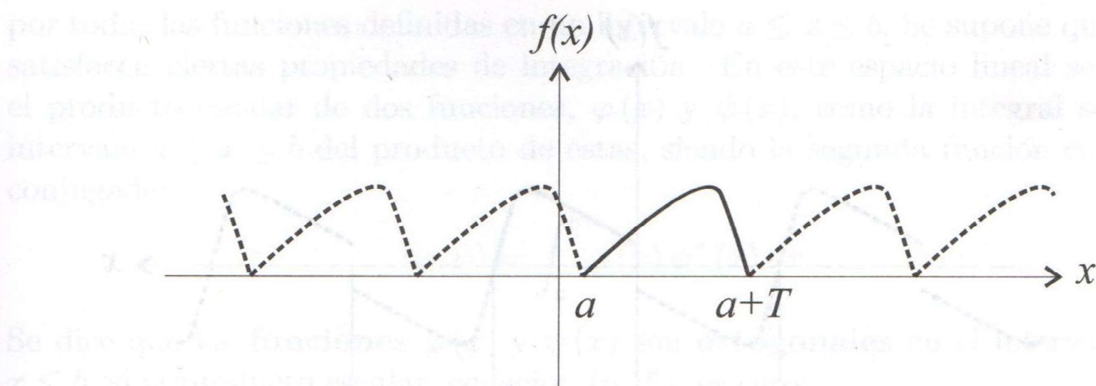



Figura 6.1. Continuación periódica de una función $\widetilde{f}(x)$ definida en $(a, a+T]$

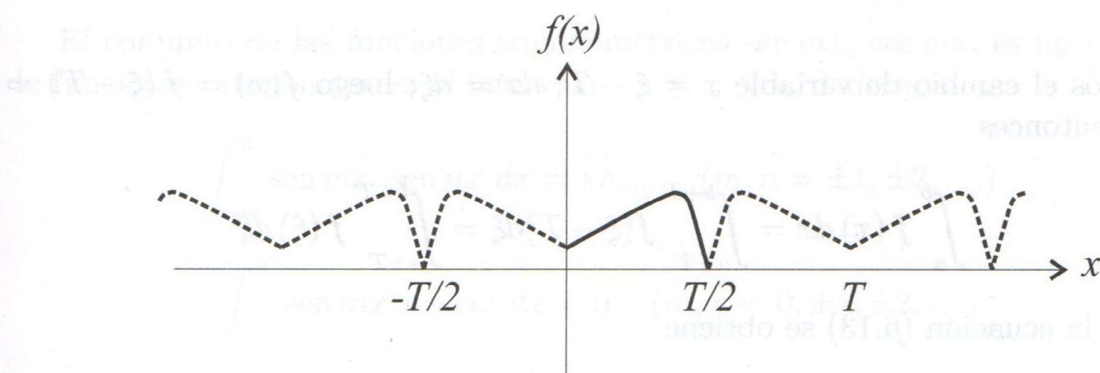



Figura 6.2. Continuación periódica par

resultando la función $\tilde{\varphi}(x)$ definida en el intervalo simétrico $-T/2 < x \leq T/2$; después,

$$\varphi(x) = \tilde{\varphi}(x - kT) \quad \text{si} \quad -T/2 + kT < x \leq T/2 + kT \tag{6.10}$$

Para obtener la función periódica impar, los pasos son como siguen (véase la figura 6.3):

$$\tilde{\chi}(x) = \begin{cases} \tilde{f}(x) & \text{si} \quad 0 < x \leq T/2 \\ -\tilde{f}(-x) & \text{si} \quad -T/2 < x \leq 0 \end{cases} \tag{6.11}$$

$$\chi(x) = \tilde{\chi}(x - kT) \quad \text{si} \quad -T/2 + kT < x \leq T/2 + kT \tag{6.12}$$

Tanto la función par $\varphi(x)$ como impar $\chi(x)$ son de periodo $T$.

Sea $f(x)$ una función periódica con periodo $T$, entonces

$$\int_a^b f(x)\,dx = \int_{a+T}^{b+T} f(x)\,dx \tag{6.13}$$

Es decir, el desplazamiento del intervalo de integración a la magnitud del periodo $T$ no cambia el valor de la integral para cualquier $a$ y $b$. Para comprobarlo,

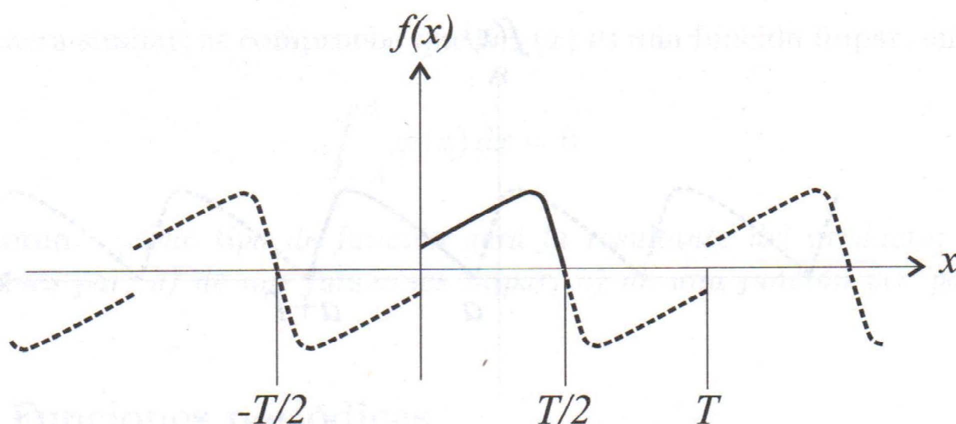



Figura 6.3. Continuación periódica impar

hacemos el cambio de variable $x = \xi - T$, $dx = d\xi$; luego $f(x) = f(\xi - T) = f(\xi)$, entonces

$$\int_a^b f(x)\,dx = \int_{a+T}^{b+T} f(\xi - T)d\xi = \int_{a+T}^{b+T} f(\xi)\,d\xi$$

De la ecuación (6.13) se obtiene

$$\int_a^{a+T} f(x)\,dx = \int_b^{b+T} f(x)\,dx \tag{6.14}$$

La comprobación es como sigue:

$$\int_a^b f(x)\,dx = \int_a^{a+T} f(x)\,dx + \int_{a+T}^{b+T} f(x)\,dx + \int_{b+T}^{b} f(x)\,dx$$

de donde

$$\int_a^b f(x)\,dx - \int_{a+T}^{b+T} f(x)\,dx - \int_{b+T}^{b} f(x)\,dx = \int_a^{a+T} f(x)\,dx$$

y, según la ecuación (6.13), se tiene

$$-\int_{b+T}^{b} f(x)\,dx = \int_a^{a+T} f(x)\,dx$$

de donde se sigue la ecuación (6.14).

### 6.1.3. Ortogonalidad de funciones

El concepto de ortogonalidad de dos vectores en un espacio lineal de dimensión finita se extiende a las funciones, considerándolas como vectores en un espacio de dimensión infinita. Con este fin, consideremos un espacio lineal constituido

por todas las funciones definidas en un intervalo $a \leq x \leq b$. Se supone que éstas satisfacen ciertas propiedades de integración. En este espacio lineal se define el producto escalar de dos funciones, $\varphi(x)$ y $\psi(x)$, como la integral sobre el intervalo $a \leq x \leq b$ del producto de éstas, siendo la segunda función complejo conjugada:

$$(\varphi, \psi) = \int_a^b \varphi(x) \psi^*(x) \, dx \tag{6.15}$$

Se dice que las **funciones** $\varphi(x)$ y $\psi(x)$ son **ortogonales** en el intervalo $a \leq x \leq b$, si el producto escalar, ecuación (6.15), es cero:

$$(\varphi, \psi) = \int_a^b \varphi(x) \psi^*(x) \, dx = 0 \tag{6.16}$$

El conjunto de las funciones trigonométricas $\operatorname{sen} mx$, $\cos mx$, es un sistema de funciones ortogonales en el intervalo $-\pi \leq x \leq \pi$, dado que

$$\int_{-\pi}^{\pi} \operatorname{sen} mx \ \operatorname{sen} nx \, dx = \pi \delta_{mn} \quad (m, n = \pm 1, \pm 2, \ldots) \tag{6.17}$$

$$\int_{-\pi}^{\pi} \operatorname{sen} mx \ \cos nx \, dx = 0 \quad (m, n = 0, \pm 1, \pm 2, \ldots) \tag{6.18}$$

$$\int_{-\pi}^{\pi} \cos mx \ \cos nx \, dx = \pi \delta_{mn} \quad (m, n = \pm 1, \pm 2, \ldots) \tag{6.19}$$

Haciendo uso de la identidad de Euler

$$e^{inx} = \cos nx + i \ \operatorname{sen} nx \tag{6.20}$$

de las fórmulas anteriores se tiene que las funciones complejas $e^{imx}$ son ortogonales sobre el intervalo $-\pi \leq x \leq \pi$:

$$\left(e^{imx}, e^{inx}\right) = \int_{-\pi}^{\pi} e^{imx} e^{-inx} dx = 2\pi \delta_{mn} \quad (m, n = 0, \pm 1, \pm 2, \ldots) \tag{6.21}$$

En la fórmula anterior se usaron las igualdades

$$(\operatorname{sen} 0x, \operatorname{sen} nx) = 0, \quad (\cos 0x, \cos nx) = 2\pi \delta_{0n} \quad (n = 0, \pm 1, \ldots)$$

## 6.2. Series de Fourier

### 6.2.1. Definición

Es fácil demostrar que la función

$$S_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n} \left(a_k \cos kx + b_k \operatorname{sen} kx\right) \tag{6.22}$$

conocida como **serie trigonométrica**, es una función periódica de periodo $T = 2\pi$, $S_n(x + 2\pi) = S_n(x)$, en donde los coeficientes $a_k$ y $b_k$ son constantes. Invirtiendo el problema, se puede preguntar si: ¿se puede aproximar una función periódica $f(x)$ de periodo $T = 2\pi$ por una serie trigonométrica $S_n(x)$? La respuesta es positiva en términos de series de Fourier.

La **serie de Fourier** de una función periódica $f(x)$ con $T = 2\pi$ es

$$Sf(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \operatorname{sen} kx) \tag{6.23}$$

o, expresando $\cos kx$ y $\operatorname{sen} kx$ en la forma exponencial mediante la fórmula de Euler, la podemos presentar en la forma exponencial

$$Sf(x) = \sum_{k=-\infty}^{\infty} \alpha_k e^{ikx} \tag{6.24}$$

donde los coeficientes de la serie se calculan según las fórmulas:

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx \, dx \quad \text{para} \quad k = 0, 1, 2, \ldots \tag{6.25}$$

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \operatorname{sen} kx \, dx \quad \text{para} \quad k = 1, 2, 3, \ldots \tag{6.26}$$

$$\alpha_k = \frac{1}{2}(a_k - ib_k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx} \, dx \quad \text{para} \quad k = 0, \pm 1, \pm 2, \ldots \tag{6.27}$$

Si $f(x)$ es real, entonces: a) $\alpha_0$ es real y $\alpha_k^* = \alpha_{-k}$; b) si $f(x)$ es real y par, entonces todos los $\alpha_k$ son reales; c) si $f(x)$ es real e impar, entonces $\alpha_0 = 0$ y los demás $\alpha_k$ son imaginarios puros.

Definiendo la suma parcial de la serie de Fourier de una función periódica $f(x)$ de periodo $2\pi$ como

$$Sf_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos kx + b_k \operatorname{sen} kx) \tag{6.28}$$

donde los coeficientes se calculan según las fórmulas anteriores; el teorema fundamental de las series de Fourier establece la convergencia de una serie de Fourier en los siguientes términos. (La demostración del teorema véala, por ejemplo, en [6] y [17])

**Teorema 6.1.** *Sea $f(x)$ una función periódica con periodo $T = 2\pi$, seccionalmente continua en el intervalo $-\pi \leq x \leq \pi$ y con derivadas continuas desde la derecha y desde la izquierda en cada punto de ese intervalo. Entonces su serie de Fourier es convergente a esta función en cada punto de continuidad, es decir,*

$$\lim_{n \to \infty} Sf_n(x) = Sf(x) = f(x)$$

*en los puntos de discontinuidad, sean $x_p$, su serie de Fourier converge a la semisuma de los límites de $f(x)$ desde la izquierda y desde la derecha del punto $x_p$, o bien,*

$$\lim_{n\to\infty} Sf_n(x_p) = Sf(x_p) = \frac{1}{2}\left[f(x_p - 0) + f(x_p + 0)\right]$$

Dicho teorema es de gran interés práctico, dado que la clase de funciones que pueden representarse mediante series de Fourier es sorprendentemente grande. Cabe mencionar que en la práctica, para la serie de Fourier correspondiente a una función $f(x)$ se usa la misma notación $f(x)$ en vez de $Sf(x)$, dado que la serie converge a la función en sus puntos de continuidad.

### 6.2.2. Propiedades

Las series de Fourier admiten la integración término a término y, además, su diferenciación, aunque la última debe hacerse con cierto cuidado, dado que puede eliminar la convergencia.

Las series de Fourier de funciones periódicas pares e impares se reducen a las **series de Fourier de cosenos** o **senos**, respectivamente.

**Teorema 6.2.** *La serie de Fourier de una función par $\varphi(x)$ de periodo $T = 2\pi$ es una serie de Fourier cosenoidal (de cosenos)*

$$\varphi(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos kx \tag{6.29}$$

*con coeficientes*

$$a_k = \frac{2}{\pi}\int_0^{\pi} \varphi(x)\cos kx\, dx \quad \textit{para} \quad k = 0, 1, 2, \ldots \tag{6.30}$$

$$b_k = 0 \quad \textit{para} \quad k = 1, 2, \ldots$$

*La serie de Fourier de una función impar $\chi(x)$ de periodo $T = 2\pi$ es una serie de Fourier senoidal (de senos)*

$$\chi(x) = \sum_{k=1}^{\infty} b_k \operatorname{sen} kx \tag{6.31}$$

*con coeficientes*

$$b_k = \frac{2}{\pi}\int_0^{\pi} \chi(x)\operatorname{sen} kx\, dx \quad \textit{para} \quad k = 1, 2, 3, \ldots \tag{6.32}$$

$$a_k = 0 \quad \textit{para} \quad k = 0, 1, 2, \ldots$$

**Demostración.** *Sea $\varphi(x)$ una función par. Entonces, los coeficientes de su serie de Fourier son*

$$a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} \varphi(x)\cos kx\ dx = \frac{2}{\pi}\int_{0}^{\pi} \varphi(x)\cos kx\ dx \quad para \quad k = 0, 1, 2, \ldots$$

$$b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} \varphi(x)\operatorname{sen} kx\ dx = 0 \quad para \quad k = 1, 2, 3, \ldots$$

*dado que el producto de dos funciones par $\varphi(x)\cos kx$, es una función par, y el producto de dos funciones $\varphi(x)\operatorname{sen} kx$, una par y otra impar, es una función impar.*

*Sea $\chi(x)$ una función impar. Entonces, coeficientes de serie de Fourier son*

$$a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} \chi(x)\cos kx\ dx = 0 \quad para \quad k = 0, 1, 2, \ldots$$

$$b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} \chi(x)\operatorname{sen} kx\ dx = \frac{2}{\pi}\int_{0}^{\pi} \chi(x)\operatorname{sen} kx\ dx \quad para \quad k = 1, 2, 3, \ldots$$

*dado que la función $\chi(x)\cos kx$ es impar, mientras tanto $\chi(x)\operatorname{sen} kx$ es par.* ■

Se pueden sumar a la clase de funciones representadas por series de Fourier las funciones definidas en intervalos finitos que, por supuesto, no son periódicas. Sea $\tilde{f}(x)$ una función definida en el intervalo $0 \leq x \leq \pi$. La podemos continuar al intervalo $-\pi \leq x \leq \pi$ obteniendo de esta manera la función $\tilde{\varphi}(x)$, ecuación (6.9), o la $\tilde{\chi}(x)$, ecuación (6.11). Luego, las funciones $\tilde{\varphi}(x)$ y $\tilde{\chi}(x)$ se continuarán periódicamente sobre todo el intervalo $-\infty \leq x \leq \infty$ dejando la función par $\varphi(x)$, ecuación (6.10), o la impar $\chi(x)$, ecuación (6.12). Las dos funciones, $\varphi(x)$ y $\chi(x)$, son periódicas con periodo $T = 2\pi$ y, además, cualquiera de ellas coincide con la función primaria $\tilde{f}(x)$ en el intervalo $0 \leq x \leq \pi$. Siendo periódicas las funciones $\varphi(x)$ y $\chi(x)$ pueden representarse mediante su serie de Fourier, cosenoidal y senoidal, respectivamente, cada una de estas series de Fourier representando la función primaria $\tilde{f}(x)$ en el intervalo $0 \leq x \leq \pi$.

**Ejemplo 6.5.** *La función $\tilde{\varphi}(x) = x^2$ definida sobre el intervalo $-\pi \leq x \leq \pi$, la podemos continuar periódicamente sobre el intervalo infinito $-\infty < x < \infty$, obteniendo la función periódica par $\varphi(x)$. Los coeficientes de la serie de Fourier de $\varphi(x)$ son:*

$$a_0 = \frac{2}{\pi}\int_{0}^{\pi} x^2\ dx = \frac{2}{3}\pi^2$$

$$a_k = \frac{2}{\pi}\int_{0}^{\pi} x^2\cos kx\ dx = (-1)^k\left(4/k^2\right) \quad para \quad k = 1, 2, \ldots$$

$$b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} x^2\operatorname{sen} kx\ dx = 0 \quad para \quad k = 1, 2, 3, \ldots$$

*La serie de Fourier de $\varphi(x)$,*

$$\varphi(x) = \frac{\pi^2}{3} + \sum_{k=1}^{\infty}(-1)^k\left(4/k^2\right)\cos kx$$

*representa la $\varphi(x)$ en todo intervalo $-\infty < x < \infty$ y reproduce la función $\tilde{\varphi}(x) = x^2$ definida sobre el intervalo $-\pi \leq x \leq \pi$. Nótese que la contribución de altas armónicas disminuye rápidamente conforme al número $k$ de la armónica.*

**Teorema 6.3.** *Sea $f(x)$ una función periódica seccionalmente continua, representada por medio de la serie de Fourier*

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \operatorname{sen} kx)$$

*Entonces, se cumple la relación*

$$(1/\pi) \int_{-\pi}^{\pi} |f(x)|^2 \, dx = \frac{a_0^2}{2} + \sum_{k=1}^{\infty} (a_k^2 + b_k^2) \tag{6.33}$$

*conocida como la **fórmula de Parseval**.*

**Demostración.** *Sustituimos $f(x)$ en forma de serie de Fourier en la integral $\int_{-\pi}^{\pi} |f(x)|^2 \, dx$. Luego, cambiamos el orden de $\int$ y $\sum$. Calculando las integrales de los productos $\cos kx \cos mx$, $\cos kx \operatorname{sen} mx$, $\operatorname{sen} kx \operatorname{sen} mx$ usaremos la ortogonalidad de las funciones armónicas, y de esta manera se obtiene la igualdad de Parseval.* ■

En física e ingeniería suelen interpretar la igualdad de Parseval en términos de la energía de una señal representada por la función $f(x)$ de la siguiente manera. La energía transmitida por la señal $f(x)$ durante un periodo $T = 2\pi$ está dada como la integral $\int_{-\pi}^{\pi} |f(x)|^2 \, dx$, de tal manera que cada armónica $\cos kx$ y $\operatorname{sen} mx$ contribuye a la energía de la señal con el peso $a_k^2$ y $b_k^2$ correspondiente.

### 6.2.3. Series de Fourier de una función de periodo arbitrario

En la práctica la mayoría de las funciones periódicas tienen sus periodos $T \neq 2\pi$. Sin embargo, estas funciones son fáciles de incluir en la clase de las representadas por medio de series de Fourier.

Sea $f(t)$ una función periódica de periodo $T$, pues $f(t+T) = f(t)$. Definimos la frecuencia angular $\omega = 2\pi/T$. Haciendo el cambio de variable $t = x/\omega$, se tiene

$$f(t) = f\left(\frac{x}{\omega}\right) = f\left(\frac{T\,x}{2\pi}\right) \equiv g(x) = g(x + 2\pi) \tag{6.34}$$

en donde la función $g(x)$ es periódica ya de periodo $2\pi$. Entonces, la podemos representar mediante la serie de Fourier

$$g(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \operatorname{sen} kx) = \sum_{k=-\infty}^{\infty} \alpha_k e^{ikx}$$

o bien,

$$f(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos k\omega t + b_k \operatorname{sen} k\omega t) = \sum_{k=-\infty}^{\infty} \alpha_k e^{ik\omega t} \qquad (6.35)$$

donde

$$a_k = \frac{2}{T} \int_{-T/2}^{T/2} f(t) \cos k\omega t \, dt \qquad (6.36)$$

$$b_k = \frac{2}{T} \int_{-T/2}^{T/2} f(t) \operatorname{sen} k\omega t \, dt \qquad (6.37)$$

$$\alpha_k = \frac{1}{2}(a_k - ib_k) = \frac{1}{T} \int_{-T/2}^{T/2} f(t) e^{-ik\omega t} \, dt \qquad (6.38)$$

Resumiendo lo anterior, vemos que la clase de funciones que pueden ser representadas mediante series de Fourier comprende todas las funciones periódicas seccionalmente continuas y, además, cualquier función definida sobre un intervalo finito, dado que ésta puede ser continuada periódicamente mediante una variedad de técnicas usando diferentes cambios de variable independiente.

Notamos que la serie de Fourier de una función periódica $f(t)$ puede considerarse como la descomposición de la función en armónicas simples $\cos k\omega t$ y $\operatorname{sen} k\omega t$ de frecuencias $\omega_k = k\omega$ que son múltiplos de la frecuencia fundamental $\omega = 2\pi/T$, puesto que los coeficientes $a_k$ y $b_k$ nos proporcionan la información sobre la contribución de cada armónica en esta descomposición. A esto se refiere cuando se habla de un análisis armónico de una señal o función; en este caso los coeficientes de la serie de Fourier, o mejor dicho, las magnitudes $\sqrt{a_k^2 + b_k^2}$ o $|\alpha_k|$, a menudo se denominan **espectro discreto de Fourier** de la función.

# 6.3. Transformada de Fourier

## 6.3.1. Definición de la transformada

La clase de funciones representadas por medio de series de Fourier es muy amplia, sin embargo, es fácil ver que todas las funciones no periódicas definidas en el intervalo infinito $-\infty < t < \infty$ están fuera de dicha clase. Un ejemplo es la función $\exp(-|x|)$. El análisis armónico de funciones de este tipo se hace por medio de la **integral** *(transformada)* **de Fourier**, que es la generalización de las series de Fourier para funciones no periódicas. Cabe mencionar que la **transformada integral** de Fourier puede introducirse considerando el límite $T \to \infty$ y $\omega = 2\pi/T \to 0$ en la ecuación (6.38). Sin embargo, la integral de Fourier la definimos de manera formal.

Para una función $f(t)$, la **transformada directa de Fourier** se define como

$$F(\omega) = \hat{F}[f(t)] \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} \, dt \qquad (6.39)$$

donde usamos la notación $\hat{F}$ [...] para el operador integral de Fourier. La **transformada inversa de Fourier** o la integral de Fourier es

$$f(t) = \hat{F}^{-1}[F(\omega)] \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} \, d\omega \tag{6.40}$$

donde $\hat{F}^{-1}$ [...] es el operador integral de la transformada inversa de Fourier. La justificación de la representación de una función por medio de una integral de Fourier se da por el siguiente teorema.

**Teorema 6.4.** *Puesto que: a) $f(t)$ es una función seccionalmente continua en todo intervalo finito y tiene la primera derivada desde la izquierda y desde la derecha en cado punto; b) $f(t)$ es absolutamente integrable sobre el intervalo $-\infty < t < \infty$, es decir, existe la integral*

$$\int_{-\infty}^{\infty} |f(t)| \, dt$$

*entonces*

$$f(t) = \hat{F}^{-1}[F(\omega)] \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} \, d\omega \tag{6.41}$$

*es decir, $f(t)$ se puede representar mediante la integral de Fourier de tal manera que en cado punto de continuidad de $f(t)$ la integral da el valor de la función y en cada punto de discontinuidad $t_p$ la integral de Fourier es igual a la semisuma de los límites de $f(t)$ desde la izquierda y desde la derecha en el punto de discontinuidad, $[f(t_p - 0) + f(t_p + 0)]/2$.*

**Ejemplo 6.6.** *Sea $f(t) = K \exp(-\alpha t^2)$ la distribución gaussiana, donde $K$ y $\alpha$ son constantes. Su transformada de Fourier es*

$$F(\omega) = \frac{K}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\alpha t^2} e^{-i\omega t} \, dt = \frac{K}{\sqrt{2\pi\alpha}} e^{-\omega^2/4\alpha} \int_{-\infty}^{\infty} e^{-u^2} \, du = \frac{K}{\sqrt{2\alpha}} e^{-\omega^2/4\alpha}$$

*en donde primero completamos el cuadrado*

$$-\alpha t^2 - i\omega t = -\left(t\sqrt{\alpha} + i\omega/2\sqrt{\alpha}\right)^2 - \omega^2/4\alpha$$

*y después cambiamos la variable de integración, $t\sqrt{\alpha} + i\omega/2\sqrt{\alpha} = u$. Nótese que el parámetro $\alpha$ determina el ancho y, por tanto, la suavidad de $f(t)$. Cuando $\alpha$ crece, el lóbulo de $f(t)$ se hace más agudo, mientras tanto $F(\omega)$ se ensancha, y viceversa. En otras palabras, cuantos más detalles finos tiene una función $f(t)$, tanto más se ensancha su transformada de Fourier $F(\omega)$, es decir, se aumenta la contribución de altas armónicas. Esta característica es genérica para la transformada de Fourier.*

**Ejemplo 6.7.** *La función pulso unitario rectangular está definida como*

$$f(t) = \begin{cases} 1 & si \ |t| \leq a \\ 0 & si \ |t| > a \end{cases}$$

*donde el parámetro es $a > 0$. La transformada de Fourier es*

$$F(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-a}^{a} e^{-i\omega t}\, dt = \frac{1}{\sqrt{2\pi}} \frac{e^{-i\omega a} - e^{i\omega a}}{-i\omega} = \sqrt{\frac{2}{\pi}} \frac{\operatorname{sen} a\omega}{\omega}$$

## 6.3.2. Propiedades de la transformada

**Dualidad.** Las fórmulas (integrales) de la transformada de Fourier directa e inversa, ecuaciones (6.39) y (6.40), se ven muy simétricas en cuanto al par de variables $(t, \omega)$ y las funciones $(f, F)$. Sustituyendo $t$ por $-\omega$ y, a la vez, $\omega$ por $t$ en la ecuación (6.40), obtenemos

$$f(-\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(t)\, e^{-i\omega t}\, dt = \hat{F}\,[F(t)] \tag{6.42}$$

es decir, $f(-\omega)$ es la transformada de Fourier de $F(t)$, lo que demuestra la dualidad del par de funciones $(f, F)$.

**Linealidad.** La transformada de Fourier es lineal, pues para dos funciones, $f_1(t)$ y $f_2(t)$, y para dos constantes arbitrarias, $C_1$ y $C_2$, de la ecuación (6.39) se tiene que

$$\hat{F}\,[C_1 f_1(t) + C_2 f_2(t)] = C_1 \hat{F}\,[f_1(t)] + C_2 \hat{F}\,[f_2(t)] = C_1 F_1(\omega) + C_2 F_2(\omega) \tag{6.43}$$

**Conjugación.** Para una función real $f(t)$ es fácil ver de la ecuación (6.39) que su transformada de Fourier $F(\omega)$ es una función compleja y, además, se tiene la propiedad de conjugación

$$F(-\omega) = F^*(\omega) \tag{6.44}$$

Dos corolarios inmediatos de esta propiedad son: a) Si $f(t)$ es par, entonces $F(\omega)$ es real; b) Si $f(t)$ es impar, entonces $F(\omega)$ es imaginaria pura.

**Problema.** *Demuestre los dos últimos corolarios.*

**Escala(miento).** Dado el par $(f(t), F(\omega))$, hallemos la transformada de Fourier de la función $f(st)$, puesto que $s$ es un número real:

$$\begin{aligned} \hat{F}\,[f(st)] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(st)\, e^{-i\omega t}\, dt = \frac{1}{s} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(u)\, e^{-i(\omega/s)u}\, du \\ &= \frac{1}{s} F\left(\frac{\omega}{s}\right) \end{aligned} \tag{6.45}$$

Esta relación llamada escalamiento, nos dice que la expansión en el tiempo, $|s| < 1$, es equivalente a la compresión en frecuencia, y viceversa.

**Desplazamiento en tiempo y frecuencia.** El cambio de la variable independiente $\tilde{t} = t - a$ en la integral de la transformada directa de Fourier conduce a la siguiente propiedad llamada **desplazamiento en tiempo**:

$$\hat{F}\,[f(t-a)] = \exp(-i\omega a)\, F(\omega) \tag{6.46}$$

donde $a$ es una constante. Similarmente, se comprueba la propiedad de **desplazamiento en frecuencia**:

$$\hat{F}\left[f(t)\,e^{i\omega_0 t}\right] = F(\omega - \omega_0) \tag{6.47}$$

**La transformada de Fourier de una derivada.** Para la transformada de Fourier de $n$-ésima derivada $f^{(n)} = \frac{d^n f}{dt^n}$ de una función $f(t)$ se cumple la igualdad

$$\hat{F}\left[f^{(n)}(t)\right] = (i\omega)^n \hat{F}\left[f(t)\right] = (i\omega)^n F(\omega) \tag{6.48}$$

que se demuestra integrando por partes en $\hat{F}\left[f^{(n)}(t)\right]$ y recordando que $f(t) \to 0$ a medida que $t \to \pm\infty$.

### 6.3.3. Teoremas de Parseval

**Teorema 6.5.** *Sean $f(t)$ y $g(t)$ dos funciones seccionalmente continuas e integrables absolutamente en el intervalo $-\infty < t < \infty$, sean $F(\omega)$ y $G(\omega)$ las transformadas de Fourier correspondientes. Entonces, son válidas*

$$\int_{-\infty}^{\infty} |F(\omega)|^2\, d\omega = \int_{-\infty}^{\infty} |f(t)|^2\, dt \tag{6.49}$$

$$\int_{-\infty}^{\infty} F(\omega)\, G(-\omega)\, d\omega = \int_{-\infty}^{\infty} f(t)\, g(t)\, dt \tag{6.50}$$

*Las dos relaciones son conocidas como el primer y segundo* ***teorema de Parseval****, respectivamente.*

**Demostración.** *Según la propiedad de conjugación*

$$G(-\omega) = G^*(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(t)\, e^{i\omega t}\, dt$$

*entonces*

$$\int_{-\infty}^{\infty} F(\omega)\, G(-\omega)\, d\omega = \int_{-\infty}^{\infty} F(\omega) \left( \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(t)\, e^{i\omega t}\, dt \right) d\omega$$

*Luego, intercambiando el orden de integración y usando la representación de Fourier de $f(t)$, ecuación (6.40), se tiene*

$$\begin{aligned} \int_{-\infty}^{\infty} F(\omega)\, G(-\omega)\, d\omega &= \int_{-\infty}^{\infty} g(t)\, dt \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\omega)\, e^{i\omega t}\, d\omega dt \\ &= \int_{-\infty}^{\infty} f(t)\, g(t)\, dt \end{aligned}$$

*lo que completa la demostración del segundo teorema de Parseval. Puesto que $g(t) = f(t)$ y, por tanto, $G(-\omega) = F(-\omega) = F^*(\omega)$, el primer teorema sigue de la relación anterior.* ■

### 6.3.4. Convolución

La **convolución** entre dos funciones, $f(t)$ y $h(t)$, está definida como

$$f * h = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(\tau) f(t-\tau)\ d\tau \tag{6.51}$$

siendo $\tau$ una variable de integración muda. Aplicando la definición se comprueba que la convolución tiene las propiedades:

$$f * h = h * f, \quad \text{es conmutativa}$$

$$f * (h_1 + h_2) = f * h_1 + f * h_2, \quad \text{es distributiva}$$

$$f * (g * h) = (f * g) * h, \quad \text{es asociativa}$$

Si $F(\omega) = \hat{F}[f(t)]$ y $H(\omega) = \hat{F}[h(t)]$, el teorema de convolución en el tiempo establece que

$$\hat{F}[f * h] = F(\omega)\, H(\omega) \tag{6.52}$$

La relación anterior dice que la transformada de Fourier de una convolución de dos funciones es el producto de las transformadas de Fourier de estas funciones, es decir, la operación de convolución en el tiempo es equivalente a la operación de multiplicación en el espacio de las transformadas de Fourier. Para probar esto calculamos

$$\begin{aligned}
\hat{F}[f * h] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left( \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(\tau) f(t-\tau)\ d\tau \right) e^{-i\omega t}\ dt \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(\tau) \left( \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t-\tau)\ e^{-i\omega t}\ dt \right) d\tau \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(\tau)\, e^{-i\omega\tau} \left( \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(u)\ e^{-i\omega u} du \right) d\tau \\
&= F(\omega) \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(\tau)\, e^{-i\omega\tau} d\tau = F(\omega)\, H(\omega)
\end{aligned}$$

lo que comprueba el teorema. En la cadena de igualdades anteriores, primero intercambiamos el orden de integración y después cambiamos la variable de integración $u = t - \tau$ y recurrimos a la definición de la transformada de Fourier.

De manera similar se demuestra la relación

$$\hat{F}^{-1}[F(\omega) * H(\omega)] = f(t)\, h(t) \tag{6.53}$$

conocida como el teorema de convolución en frecuencia.

Además, tenemos la inversa de las relaciones anteriores:

$$f * h = \hat{F}^{-1}[F(\omega)\ H(\omega)]$$

$$F(\omega) * H(\omega) = \hat{F}[f(t)\, h(t)]$$

donde $\hat{F}^{-1}$ denota la transformada inversa de Fourier.

**Nota bibliográfica:** para ampliar los conocimientos sobre el tema de este capítulo, puede consultar [1], [4], [6], [12] y [17].

# Capítulo 7

# Transformada de Laplace

Con frecuencia en la física matemática se encuentran pares de funciones relacionadas mediante una expresión de la forma

$$F(\alpha) = \int_a^b f(t)\, N(\alpha, t)\, dt \tag{7.1}$$

La función $F(\alpha)$ se denomina **transformada integral** de $f(t)$ por el núcleo $N(\alpha, t)$. Dentro de un mundo de transformadas integrales, las transformadas más conocidas son la de Fourier (estudiada en el capítulo 6)

$$F(\omega) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} f(t)\, e^{-i\omega t} dt \tag{7.2}$$

con el núcleo $N(\omega, t) = (2\pi)^{-1/2} e^{-i\omega t}$, la **transformada de Laplace unilateral**

$$F(\alpha) = \int_0^{\infty} f(t)\, e^{-\alpha t} dt, \quad N(\alpha, t) = e^{-\alpha t} \tag{7.3}$$

la **transformada de Hankel (Fourier-Bessel)**

$$F(\alpha) = \int_0^{\infty} f(t)\, t J_n(\alpha t)\, dt, \quad N(\alpha, t) = t J_n(\alpha t) \tag{7.4}$$

en donde $J_n(t)$ es la función de Bessel de orden $n$, y la **transformada de Mellin**

$$F(\alpha) = \int_0^{\infty} f(t)\, t^{\alpha-1} dt, \quad N(\alpha, t) = t^{\alpha-1} \tag{7.5}$$

Dichas transformadas son de utilidad en el análisis matemático y en las aplicaciones a la física e ingeniería. Cabe mencionar que todas estas transformadas integrales son lineales; es decir, denotando cualquiera de estas transformadas por $L\{f(t)\}$, se tiene

$$L\{c_1 f_1(t) + c_2 f_2(t)\} = c_1 L\{f_1(t)\} + c_2 L\{f_2(t)\} \tag{7.6}$$

En este capítulo se considera la transformada de Laplace desde un punto de vista práctico como un método para resolver ecuaciones diferenciales y los problemas correspondientes con valores en la frontera e iniciales. La aplicación del método a las ecuaciones diferenciales parciales se ilustra en el capítulo 8. El método se utiliza ampliamente en las matemáticas de ingeniería y física. Resulta particularmente útil para las ecuaciones diferenciales lineales no homogéneas en las que la fuerza motriz tiene discontinuidades. Además, las ecuaciones no homogéneas se pueden resolver sin encontrar primero la solución a la ecuación homogénea correspondiente. Otra ventaja es que los problemas con valores iniciales se pueden resolver sin determinar primero la solución general.

## 7.1. Definición y propiedades básicas

Sea $f(t)$ una función dada que está definida para toda $-\infty < t < \infty$. Si la integral

$$F(z) = \int_{-\infty}^{\infty} f(t)\, e^{-zt} dt \tag{7.7}$$

existe cuando menos para una recta $\operatorname{Re} z = s$, entonces la función $F(z)$ se llama **transformada de Laplace bilateral** de $f(t)$.

Si la función $f(t)$ está definida para toda $t \geq 0$ y la integral

$$F(z) = \int_{0}^{\infty} f(t)\, e^{-zt} dt \tag{7.8}$$

existe en un semiplano $\operatorname{Re} z > c$, entonces la función $F(z)$ se llama **transformada de Laplace unilateral** de $f(t)$ y se denota $L\{f(t)\}$. La transformada de Laplace unilateral puede considerarse como un caso particular de la bilateral de una función que es cero para $t < 0$. En la práctica, es más usada la transformada unilateral y, además, frecuentemente la variable $z = s$ es real (véase la ecuación (7.3)).

La función $f(t)$ que aparece en la definición (7.8) se conoce como (transformada) **inversa** de $F(z)$ y se denota por

$$f(t) = L^{-1}\{F(z)\} \tag{7.9}$$

La unicidad y otros aspectos de la correspondencia $F(z) \div f(t)$ se consideran en la sección 7.4.

Nótese que la transformada de Laplace bilateral, $F(z) = \hat{L}\{f(t)\}$, de una función $f(t)$ está estrechamente relacionada con la transformada de Fourier, $\mathcal{F}(\omega) = \hat{F}\{f(t)\}$, de la misma función, es decir, $F(i\omega) = \mathcal{F}(\omega)$, en donde $-\infty < \omega < \infty$. De esta relación, las propiedades de la transformada de Laplace se obtienen fácilmente de las de la transformada de Fourier, y viceversa. Las propiedades de la transformada de Laplace son como siguen.

Ante todo, la existencia de las transformadas de Laplace se da por el siguiente teorema.

**Teorema 7.1.** *Sea $f(t)$ una función seccionalmente continua sobre el intervalo $t \geq 0$, que satisface la condición*

$$|f(t)| \leq M e^{\gamma t} \tag{7.10}$$

*para algunas constantes $M$ y $\gamma$. Entonces, la transformada de Laplace de $f(t)$ existe para toda $s > \gamma$.*

**Demostración.** *Como $f(t)$ es seccionalmente continua, $e^{-st}f(t)$ es integrable sobre cualquier intervalo finito del eje $t$. Suponiendo que $s > \gamma$, de (7.10) se obtiene*

$$\left|\int_0^\infty f(t)\, e^{-st} dt\right| \leq \int_0^\infty |f(t)|\, e^{-st} dt \leq M \int_0^\infty e^{\gamma t} e^{-st} dt = \frac{M}{s-\gamma}$$

*en donde se necesitó la condición $s > \gamma$ para la existencia de la última integral. Esto completa la demostración.* ■

Las condiciones del teorema son suficientes para la mayor parte de las aplicaciones y son fáciles para averiguar si una función las satisface. Por ejemplo, cualquier función cuyo valor absoluto sea acotado para toda $t \geq 0$, como las funciones $\operatorname{sen} t$ y $\cos t$, satisface las condiciones del teorema. Otros ejemplos son

$$\cosh t < e^t, \quad t^n < n! e^t \quad (n = 0, 1, 2, \ldots), \quad t \geq 0$$

Tenga en cuenta que las condiciones del teorema son suficientes más que necesarias. Por lo tanto, por ejemplo, la función $f(t) = t^{-1/2}$ es infinita en $t = 0$, sin embargo, partiendo de la definición de función gamma y de que $\Gamma(1/2) = \sqrt{\pi}$, es fácil ver que su transformada existe:

$$\begin{aligned} L\left\{t^{-1/2}\right\} &= \int_0^\infty t^{-1/2} e^{-st} dt = s^{-1/2} \int_0^\infty x^{-1/2} e^{-x} dx \\ &= s^{-1/2} \Gamma(1/2) = (\pi/s)^{1/2} \end{aligned}$$

La **linealidad** de la transformada de Laplace es dada por la fórmula (7.6). Algunos ejemplos de las transformadas de Laplace son como siguen. En todos los casos se supone que $f(t) = 0$ para $t < 0$.

**Ejemplo 7.1.** *Sea $f(t) = 1$ para $t \geq 0$. Entonces*

$$L\{1\} = \int_0^\infty e^{-st} dt = 1/s, \quad s > 0 \tag{7.11}$$

**Ejemplo 7.2.** *Sea $f(t) = e^{at}$ para $t \geq 0$. En este caso*

$$L\{e^{at}\} = \int_0^\infty e^{at} e^{-st} dt = \frac{1}{s-a}, \quad s > a \tag{7.12}$$

**Ejemplo 7.3.** *Partiendo de la linealidad de la transformada de Laplace y de la ecuación (7.12), se obtienen*

$$L\{\cosh at\} = \frac{1}{2}\left(L\{e^{at}\} + L\{e^{-at}\}\right) = \frac{1}{2}\left(\frac{1}{s-a} + \frac{1}{s+a}\right) = \frac{s}{s^2-a^2} \quad (7.13)$$

$$L\{\operatorname{senh} at\} = \frac{1}{2}\left(L\{e^{at}\} - L\{e^{-at}\}\right) = \frac{1}{2}\left(\frac{1}{s-a} - \frac{1}{s+a}\right) = \frac{a}{s^2-a^2} \quad (7.14)$$

*ambas válidas para $s > a$. Mediante las relaciones*

$$\cos at = \cosh iat, \quad \operatorname{sen} t = -i \operatorname{senh} iat$$

*se obtienen las transformadas de Laplace*

$$L\{\cos at\} = \frac{s}{s^2+a^2}, \quad L\{\operatorname{sen} at\} = \frac{a}{s^2+a^2} \quad (7.15)$$

**Ejemplo 7.4.** *Para $f(t) = t^a$, se tiene*

$$L\{t^a\} = \int_0^\infty t^a e^{-st} dt = \frac{1}{s^{a+1}} \int_0^\infty x^a e^{-x} dx = \frac{1}{s^{a+1}} \Gamma(a+1) \quad (7.16)$$

*para $s > 0$ y $a > -1$.*

**Desplazamiento sobre el eje $s$.** Si $F(s) = L\{f(t)\}$ para $s > \gamma$, entonces $F(s-a) = L\{e^{at} f(t)\}$, en donde $s - a > \gamma$. Es decir, la sustitución de $s$ por $s - a$ en la transformada (desplazamiento sobre el eje $s$) corresponde a la multiplicación de la función original $f(t)$ por $e^{at}$. La demostración es fácil. Por definición,

$$F(s) = \int_0^\infty f(t) e^{-st} dt$$

y, por tanto,

$$F(s-a) = \int_0^\infty f(t) e^{-(s-a)t} dt = \int_0^\infty e^{at} f(t) e^{-st} dt = L\{e^{at} f(t)\} \quad (7.17)$$

**Ejemplo 7.5.** *De las ecuaciones (7.15), (7.16) y (7.17) se tiene*

$$L\{e^{at} \cos \omega t\} = \frac{s-a}{(s-a)^2 + \omega^2}, \quad L\{e^{at} \operatorname{sen} \omega t\} = \frac{\omega}{(s-a)^2 + \omega^2} \quad (7.18)$$

$$L\{e^{at} t^n\} = \frac{1}{(s-a)^{n+1}} \Gamma(n+1) \quad (7.19)$$

**Desplazamiento sobre el eje $t$.** Si $F(s) = L\{f(t)\}$ es la transformada de $f(t)$, entonces $e^{-as} F(s) = L\left\{\tilde{f}(t)\right\}$ es la transformada de la función

$$\tilde{f}(t) = \begin{cases} f(t-a) & \text{si} \quad t \geq a \\ 0 & \text{si} \quad t < a \end{cases} \quad (7.20)$$

Nótese que mediante la **función escalón unitario** $u_a(t)$, que es

$$u_a(t) = \begin{cases} 1 & \text{si} \quad t \geq a \\ 0 & \text{si} \quad t < a \end{cases} \tag{7.21}$$

la función $\tilde{f}(t)$ puede escribirse como $\tilde{f}(t) = f(t-a)\,u_a(t)$. Por tanto, el desplazamiento sobre el eje $t$, afirma que

$$e^{-as}F(s) = L\{f(t-a)\,u_a(t)\} \tag{7.22}$$

La demostración es como sigue. Por definición, se tiene

$$\begin{aligned} e^{-as}F(s) &= \int_0^{\infty} f(\tau)\,e^{-s(\tau+a)}d\tau = \int_a^{\infty} f(t-a)\,e^{-st}dt \\ &= \int_0^{\infty} f(t-a)\,u_a(t)\,e^{-st}dt = L\{f(t-a)\,u_a(t)\} \end{aligned}$$

Quedando completa la demostración.

La función escalón unidad es muy importante, dado que es la pieza básica de ciertas funciones de gran utilidad. Por lo tanto, conviene señalar que la transformada de Laplace de $u_a(t)$ es

$$L\{u_a(t)\} = \frac{e^{-as}}{s} \tag{7.23}$$

que se deduce directamente de la definición o aplicando la fórmula (7.22) a la función unidad $f(t) = 1$.

## 7.2. Transformada de derivadas e integrales

Una propiedad de la transformada de Laplace importante para el cálculo operacional (véase la sección 7.4) es la transformada de la derivada y de la integral de una función $f(t)$. Dada cierta similitud entre las transformadas de Fourier y de Laplace, es de esperar que la derivación de una función $f(t)$ corresponda simplemente a la multiplicación de la transformada $F(s)$ por $s$. Además, la integración de $f(t)$, siendo la operación inversa de la derivación, es de esperar que corresponda a la división de la transformada entre $s$. En efecto, salvo que la transformada de Laplace, siendo una transformada unilateral, involucra el valor inicial de la función en la fórmula de transformada de la derivada.

### 7.2.1. Transformada de la derivada

Procedamos a transformar la primera derivada de $f(t)$. En la práctica suelen encontrarse con los dos casos que se consideran a continuación. En todo caso se supone que $f(t)$, para algunas constantes $M$ y $\gamma$, satisface la condición

$$|f(t)| \leq Me^{\gamma t} \tag{7.24}$$

**Caso 1**. La función $f(t)$ es continua y su primera derivada $df(t)/dt$ es seccionalmente continua en el intervalo $t \geq 0$, siendo $0 = a_0 < a_1 < ... < a_{K+1} = \infty$ los puntos de discontinuidad de la derivada. En este caso el intervalo de integración se divide en partes tales que $df(t)/dt$ sea continua en cada una de ellas. Por definición y al integrar por partes, se tiene

$$\begin{aligned} L\{df(t)/dt\} &= \int_0^\infty e^{-st}\frac{df(t)}{dt}dt = \sum_{k=0}^{K}\left(\int_{a_k}^{a_{k+1}} e^{-st}\frac{df(t)}{dt}dt\right) \\ &= \sum_{k=0}^{K}\left(e^{-st}f(t)\Big|_{a_k}^{a_{k+1}} + s\int_{a_k}^{a_{k+1}} e^{-st}f(t)\,dt\right) \\ &= -f(0) + s\int_0^\infty e^{-st}f(t)\,dt \end{aligned}$$

cuando $s > \gamma$. Entonces,

$$L\{df(t)/dt\} = sF(s) - f(0) \tag{7.25}$$

Por supuesto, la fórmula es también válida para la derivada $df(t)/dt$ continua en el intervalo $t \geq 0$.

**Caso 2**. La función $f(t)$ es seccionalmente continua y su primera derivada $df(t)/dt$ continua en el intervalo $t \geq 0$, siendo $0 = d_0 < d_1 < ... < d_{M+1} = \infty$ los puntos de discontinuidad de la función $f(t)$. En este caso el intervalo de integración se divide en partes tales que la función $f(t)$ sea continua en cada una de ellas. Al integrar por partes, se tiene

$$\begin{aligned} L\{df(t)/dt\} &= \int_0^\infty e^{-st}\frac{df(t)}{dt}dt = \sum_{m=0}^{M}\left(\int_{d_m}^{d_{m+1}} e^{-st}\frac{df(t)}{dt}dt\right) \\ &= \sum_{m=0}^{M}\left(e^{-st}f(t)\Big|_{d_m+0}^{d_{m+1}-0} + s\int_{d_m}^{d_{m+1}} e^{-st}f(t)\,dt\right) \\ &= -f(0) + \sum_{m=1}^{M} e^{-sd_m}\left[f(d_m-0) - f(d_m+0)\right] \\ &\quad + s\int_0^\infty e^{-st}f(t)\,dt \end{aligned}$$

para $s > \gamma$. Entonces,

$$L\{df(t)/dt\} = sF(s) - f(0) - \sum_{m=1}^{M} e^{-sd_m}\left[f(d_m+0) - f(d_m-0)\right] \tag{7.26}$$

**Caso 3**. Tanto la función $f(t)$ como su primera derivada $df(t)/dt$ son seccionalmente continuas en el intervalo $t \geq 0$, siendo $0 = d_0 < d_1 < \cdots < d_{M+1} = \infty$ los puntos de discontinuidad de $f(t)$ y $0 = a_0 < a_1 < \cdots <$

$a_{K+1} = \infty$ los puntos de discontinuidad de la derivada. Nótese que algunos de los puntos $\{d_m\}$ pueden coincidir con los de $\{a_k\}$. Para calcular la transformada de la derivada, el intervalo de integración se divide en partes tales que tanto la función $f(t)$ como su primera derivada $df(t)/dt$ sean continuas en cada una de ellas. Al integrar por partes en cada uno de dichos subintervalos, se concluye que los puntos de discontinuidad de la derivada, $\{a_k\}$, no alteran la fórmula de la transformada, pero los puntos $\{d_m\}$ contribuyen a dicha fórmula, quedando finalmente validada la fórmula (7.26).

Al aplicar (7.25) a la segunda derivada $d^2 f(t)/dt^2$, se obtiene

$$\begin{aligned} L\left\{d^2 f(t)/dt^2\right\} &= sL\{df(t)/dt\} - df(0)/dt \\ &= s[sF(s) - f(0)] - df(0)/dt \\ &= s^2 F(s) - sf(0) - f^{(1)}(0) \end{aligned} \tag{7.27}$$

De manera similar, por inducción se demuestra que, para $s > \gamma$, la transformada de Laplace de $n$-ésima derivada $f^{(n)}(t)$ es

$$L\left\{f^{(n)}\right\} = s^n F(s) - s^{n-1} f(0) - s^{n-2} f^{(1)}(0) - \cdots - f^{(n-1)}(0) \tag{7.28}$$

suponiendo que $f(t)$ y sus derivadas $f^{(1)}(t), f^{(2)}(t), \ldots, f^{(n-1)}(t)$ son continuas para $t \geq 0$ y satisfacen la condición $\left|f^{(k)}(t)\right| \leq Me^{\gamma t}$ para algunas constantes $M$ y $\gamma$, y la derivada $f^{(n)}(t)$ es seccionalmente continua sobre el intervalo $t \geq 0$.

**Ejemplo 7.6.** *Sea $f(t) = t \operatorname{sen} at$. Para hallar $L\{t \operatorname{sen} at\}$, notemos que $f(0) = 0$ y, además,*

$$\begin{aligned} f^{(1)}(t) &= \operatorname{sen} at + at \cos at, \quad f^{(1)}(0) = 0 \\ f^{(2)}(t) &= 2a \cos at - a^2 f(t), \end{aligned}$$

*de modo que, aplicando (7.28) a la última relación, se tiene*

$$s^2 F(s) = 2aL\{\cos at\} - a^2 F(s)$$

*Finalmente, aplicando la fórmula de la transformada de $\cos at$, ecuación (7.15), se obtiene*

$$L\{t \operatorname{sen} at\} = F(s) = \frac{2as}{\left(s^2 + a^2\right)^2}$$

### 7.2.2. Transformada de la integral

Dado que la integración de $f(t)$ es la operación inversa de la derivación, es de esperar que la transformada de la integral corresponda a la división entre $s$. En efecto, si $f(t)$ es seccionalmente continua y satisface la condición (7.24), entonces

$$L\left\{\int_0^t f(\tau)\, d\tau\right\} = \frac{1}{s} L\{f(t)\}, \quad (s > \gamma > 0) \tag{7.29}$$

Para demostrar la fórmula, notemos que la función

$$g(t) = \int_0^t f(\tau)\, d\tau$$

es continua. Al aplicar (7.24) se obtiene

$$|g(t)| \leq \int_0^t |f(\tau)|\, d\tau \leq M \int_0^t e^{\gamma\tau} d\tau = \frac{M}{\gamma}\left(e^{\gamma t} - 1\right)$$

es decir, la función $g(t)$ satisface una condición del tipo (7.24). Además, $dg(t)/dt = f(t)$ es seccionalmente continua para $t \geq 0$. Por tanto, usando (7.25), se tiene

$$L\{f(t)\} = L\left\{g^{(1)}(t)\right\} = sL\{g(t)\} - g(0) = sL\{g(t)\}$$

y esto completa la demostración de (7.29).

Al tomar la transformada inversa de ambos términos de la igualdad (7.29), se obtiene la fórmula útil

$$L^{-1}\left\{\frac{1}{s}F(s)\right\} = \int_0^t f(\tau)\, d\tau \tag{7.30}$$

en donde $F(s) = L\{f(t)\}$. En ocasiones esta fórmula permite recuperar la transformada inversa de Laplace, es decir, la función $f(t)$, como se explica a continuación.

**Ejemplo 7.7.** *Sea $L\{f(t)\} = a/\left(s^2\left(s^2 + a^2\right)\right)$. Con el fin de encontrar $f(t)$, presentamos la transformada en la forma*

$$L\{f(t)\} = \frac{1}{s}\left(\frac{1}{s}\frac{a}{s^2 + a^2}\right) = \frac{1}{s}\left(\frac{1}{s}L\{\operatorname{sen} at\}\right) = \frac{1}{s}\tilde{F}(s)$$

*donde*

$$\tilde{F}(s) = \frac{1}{s}L\{\operatorname{sen} at\}, \qquad \frac{a}{s^2 + a^2} = L\{\operatorname{sen} at\}$$

*Al aplicar (7.30) se tiene*

$$f(t) = L^{-1}\left\{\frac{1}{s}\tilde{F}(s)\right\} = \int_0^t \tilde{f}(\tau)\, d\tau$$

*donde*

$$\begin{aligned} \tilde{f}(t) &= L^{-1}\left\{\frac{1}{s}L\{\operatorname{sen} at\}\right\} = \int_0^t \operatorname{sen} a\tau d\tau \\ &= \frac{1}{a}(1 - \cos at) \end{aligned}$$

*Finalmente,*

$$f(t) = \int_0^t \tilde{f}(\tau)\, d\tau = \frac{1}{a}\left(t - \frac{\operatorname{sen} at}{a}\right)$$

## 7.3. Derivación e integración de la transformada. Convolución

Como hemos visto, la transformada de Laplace tiene varias propiedades generales que pueden usarse para obtener tanto transformadas como sus inversas. Los métodos basados en esas propiedades son: la integración directa, la aplicación de la linealidad, traslación sobre el eje $s$ o $t$, derivación e integración de las funciones originales $f(t)$. Estos métodos se completan con la derivación e integración de las transformadas y se encuentran las operaciones correspondientes en las funciones originales $f(t)$. Además, se considerará otra propiedad importante que tiene que ver con el producto de transformadas; la operación correspondiente en las funciones originales se llama convolución.

### 7.3.1. Derivación e integración de la transformada

Se puede demostrar que si $f(t)$ satisface las condiciones del teorema de existencia dado en la sección 7.1, entonces se puede obtener la derivada de la transformada correspondiente

$$F(s) = L\{f(t)\} = \int_0^\infty f(t)\, e^{-st} dt$$

con respecto a $s$, derivando bajo el signo de la integral, quedando

$$F'(s) = -\int_0^\infty f(t)\, te^{-st} dt = -L\{tf(t)\} \quad \text{o} \quad L\{tf(t)\} = -F'(s) \qquad (7.31)$$

Entonces, la derivación de la transformada de una función corresponde a la multiplicación por $-t$. Esta propiedad hace posible obtener nuevas transformadas a partir de otras dadas.

**Ejemplo 7.8.** *Aplicando la linealidad de la transformada y la ecuación (7.31), a partir de las transformadas $L\{\operatorname{sen} at\}$ y $L\{\cos at\}$, ecuación (7.15), es fácil obtener las siguientes transformadas:*

$$L\{t \operatorname{sen} at\} = \frac{2as}{(s^2+a^2)^2}$$

$$L\{t\cos at\} = \frac{2s^2}{(s^2+a^2)^2} - \frac{1}{s^2+a^2} = \frac{s^2-a^2}{(s^2+a^2)^2}$$

$$L\left\{\frac{1}{2a}(\operatorname{sen} at + at\cos at)\right\} = \frac{1}{2(s^2+a^2)} + \frac{s^2-a^2}{2(s^2+a^2)^2} = \frac{s^2}{(s^2+a^2)^2}$$

$$L\left\{\frac{1}{2a^3}(\operatorname{sen} at - at\cos at)\right\} = \frac{1}{2a^2(s^2+a^2)} - \frac{s^2-a^2}{2a^2(s^2+a^2)^2}$$
$$= \frac{1}{(s^2+a^2)^2}$$

**Ejemplo 7.9.** *Con el fin de encontrar la función (transformada inversa) $f(t)$ que corresponde a la transformada $L\{f(t)\} = \ln\left(s^2/\left(s^2+a^2\right)\right) = F(s)$, se puede proceder de la siguiente manera. Al derivar la transformada se tiene*

$$F'(s) = \frac{2a^2}{s(s^2+a^2)} = \frac{2}{s} - \frac{2s}{s^2+a^2}$$

*Dado que $L\{tf(t)\} = -F'(s)$, se obtiene $tf(t) = -L^{-1}\{F'(s)\} = -2 + 2\cos at$. Entonces,*

$$f(t) = L^{-1}\left\{\ln\left(s^2/\left(s^2+a^2\right)\right)\right\} = 2(\cos at - 1)/t$$

De manera similar, si $f(t)$ satisface las condiciones del teorema de existencia dado en la sección 7.1 y $\lim_{t\to+0}(f(t)/t)$ existe, entonces

$$L\left\{\frac{f(t)}{t}\right\} = \int_s^\infty F(u)\,du, \quad \text{para } s > \gamma \tag{7.32}$$

es decir, la integración de la transformada de una función $f(t)$ corresponde a la división de $f(t)$ entre $t$. En efecto, cuando $s > \gamma$,

$$\int_s^\infty F(u)\,du = \int_s^\infty \left(\int_0^\infty f(t)\,e^{-ut}dt\right)du = \int_0^\infty f(t)\left(\int_s^\infty e^{-ut}du\right)dt$$
$$= \int_0^\infty \frac{f(t)}{t}e^{-st}dt = L\left\{\frac{f(t)}{t}\right\}$$

en donde, bajo las suposiciones anteriores, es posible invertir el orden de la integración.

### 7.3.2. Convolución

A menudo sucede que se dan dos transformadas, $F(s)$ y $G(s)$, cuyas inversas, $f(t)$ y $g(t)$, se conocen y se necesita calcular la inversa del producto $H(s) = F(s)G(s)$. Resulta que la inversa $h(t) = L^{-1}\{H(s)\}$ está relacionada con las inversas $f(t)$ y $g(t)$ a través de la operación de integración conocida como **convolución** (unilateral) de $f(t)$ y $g(t)$. La convolución unilateral la denotamos por $(f * g)$, y se define como

$$(f * g) = \int_0^t f(\tau)\,g(t-\tau)\,d\tau \tag{7.33}$$

La convolución unilateral puede considerarse como un caso particular de la bilateral, definida en la sección 6.3, para las funciones $f(t)$ y $g(t)$ definidas

para $t > 0$, tales que $f(t) = 0$ y $g(t) = 0$ para $t < 0$. Esta observación hace válidas las siguientes propiedades de la convolución:

$$(f * g) = (g * f), \quad \text{es conmutativa}$$

$$(f * (g_1 + g_2)) = (f * g_1 + f * g_2), \quad \text{es distributiva}$$

$$(f * (g * k)) = ((f * g) * k), \quad \text{es asociativa}$$

Lo dicho anteriormente se resume en el siguiente teorema.

**Teorema 7.2.** *Si $f(t)$ y $g(t)$ son las transformadas inversas de $F(s)$ y $G(s)$ respectivamente, y satisfacen las hipótesis del teorema de existencia de la transformada de Laplace, entonces, la transformada inversa $h(t)$ del producto $H(s) = F(s)G(s)$ es la convolución de $f(t)$ y $g(t)$,*

$$L\{(f * g)\} = F(s)G(s) \quad o \tag{7.34}$$

$$L^{-1}\{F(s)G(s)\} = h(t) = (f * g) = \int_0^t f(\tau)\, g(t-\tau)\, d\tau \tag{7.35}$$

**Demostración.** *Por la definición de $G(s)$ y la propiedad de desplazamiento sobre el eje $t$, ecuación (7.22), para $\tau \geq 0$ fija, se tiene*

$$e^{-s\tau} G(s) = \int_0^\infty g(t-\tau)\, u_\tau(t)\, e^{-st} dt$$

*en donde $s > \gamma$. A partir de esto y de la definición de $F(s)$, se obtiene*

$$\begin{aligned} F(s)G(s) &= \int_0^\infty f(\tau)\, e^{-s\tau} G(s)\, d\tau \\ &= \int_0^\infty f(\tau) \int_0^\infty g(t-\tau)\, u_\tau(t)\, e^{-st} dt d\tau \\ &= \int_0^\infty \left( \int_0^\infty f(\tau)\, g(t-\tau)\, u_\tau(t)\, d\tau \right) e^{-st} dt \\ &= \int_0^\infty \left( \int_0^t f(\tau)\, g(t-\tau)\, d\tau \right) e^{-st} dt \\ &= L\{(f * g)\} \end{aligned}$$

*que completa la demostración del teorema.* ■

Como se verá adelante, la convolución es útil para obtener transformadas inversas y, por lo tanto, para resolver ecuaciones diferenciales.

**Ejemplo 7.10.** *Sea $H(s) = 1/\left(s^2(s-a)\right)$. Ésta se factoriza naturalmente en el producto de $F(s) = 1/s^2$ y $G(s) = 1/(s-a)$ que tienen las inversas*

$$f(t) = L^{-1}\{F(s)\} = t, \quad g(t) = L^{-1}\{G(s)\} = e^{at}$$

*Por tanto, la inversa* $h(t)$ *de* $H(s)$ *es*

$$\begin{aligned} h(t) &= L^{-1}\{H(s)\} = \int_0^t f(\tau)\, g(t-\tau)\, d\tau \\ &= \int_0^t \tau e^{a(t-\tau)} d\tau = \frac{1}{a^2}\left(e^{at} - at - 1\right) \end{aligned}$$

## 7.4. Cálculo operacional

El proceso de resolución de una ecuación diferencial ordinaria consta principalmente de tres pasos:

1) La ecuación diferencial dada se transforma en una ecuación subsidiaria que, en el caso de ecuaciones diferenciales ordinarias, es algebraica y más sencillo resolver.

2) La ecuación subsidiaria se resuelve a través de operaciones puramente algebraicas.

3) A la solución de la ecuación subsidiaria se le aplica la transformada inversa con el fin de obtener la solución de la ecuación diferencial dada.

De esta manera, el problema de resolver una ecuación diferencial se reduce al de solucionar un problema algebraico. El primer paso se hace mediante las tablas de correspondencia entre las funciones y sus transformadas y/o aplicando las propiedades conocidas de transformadas. El segundo paso es puramente algebraico. Por lo general, los dos primeros pasos se realizan sin problemas. El tercer paso requiere la transformación inversa de una expresión algebraica, por lo general, bastante compleja. Este paso se efectúa mediante las tablas y/o aplicando las propiedades de transformadas inversas. El método basado en dichos pasos recibe el nombre de cálculo operacional.

El cálculo operacional se utiliza ampliamente en las matemáticas de ingeniería, en donde tiene numerosas aplicaciones. Resulta particularmente útil para resolver ecuaciones diferenciales ordinarias con coeficientes constantes y con valores iniciales. Otra ventaja es que las ecuaciones no homogéneas se pueden resolver sin encontrar primero la solución general a la ecuación homogénea correspondiente. El método de la transformada de Fourier también se extiende a las ecuaciones diferenciales ordinarias con coeficientes variables, ecuaciones diferenciales parciales y las integrales.

### 7.4.1. Ecuaciones diferenciales ordinarias con coeficientes constantes

Considérese la ecuación diferencial ordinaria de orden $n$ con coeficientes constantes $a_i$ $(i = 1, 2, \ldots, n)$

$$y^{(n)} + a_1 y^{(n-1)} + a_2 y^{(n-2)} + \cdots + a_{n-1} y^{(1)} + a_n y = f(t) \tag{7.36}$$

en donde $f(t)$ es una función conocida para $t > 0$. Con el fin de simplificar las expresiones, suponemos que tanto la solución como sus derivadas hasta el orden $n$ son funciones continuas en todo el intervalo $t > 0$.

Dado que la transformada de Laplace de la derivada de una función depende del valor de esta función en $t = 0$, primero conviene tratar el caso cuando las condiciones iniciales son cero

$$y(0) = 0,\ y^{(1)}(0) = 0,\ \ldots,\ y^{(n-1)}(0) = 0 \tag{7.37}$$

Al aplicar la transformada de Laplace a la ecuación (7.36) bajo la condición (7.37), mediante las fórmulas (7.25) y (7.28), se obtiene la ecuación subsidiaria

$$s^n Ly + a_1 s^{n-1} Ly + a_2 s^{n-2} Ly + \cdots + a_{n-1} s Ly + a_n Ly = Lf(t)$$

En forma compacta, ésta es

$$P(s)\, Y(s) = F(s) \tag{7.38}$$

en donde $Y(s) = Ly$, $F(s) = Lf$ y $P(s) = s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n$ es polinomio característico de (7.36).

No es difícil tomar en cuenta las condiciones iniciales generales

$$y(0) = y_0,\ y^{(1)}(0) = y_1,\ \ldots,\ y^{(n-1)}(0) = y_{n-1} \tag{7.39}$$

Según la fórmula (7.28), cada término $Ly^{(m)}$ $(m = 1, 2, \ldots, n)$ contiene un polinomio adicional de orden $(m - 1)$ en $s$ que no contiene la transformada $Y(s)$. Por tanto, la aplicación de la transformada a la ecuación (7.36) bajo la condición (7.37) lleva a la ecuación subsidiaria

$$P(s)\, Y(s) - P_0(s) = F(s) \tag{7.40}$$

en donde el polinomio $P_0(s)$ de orden $(n - 1)$ depende de las condiciones iniciales,

$$\begin{aligned} P_0(s) = \ & s^{n-1} y_0 + s^{n-2}(y_1 + a_1 y_0) + \cdots + \\ & (y_{n-1} + a_1 y_{n-2} + \cdots + a_{n-2} y_1 + a_{n-1} y_0) \end{aligned}$$

La relación (7.40) conduce a la fórmula importante

$$Y(s) = \frac{F(s)}{P(s)} + \frac{P_0(s)}{P(s)} \tag{7.41}$$

para la transformada de la solución $Y(s)$. Aplicando la transformada inversa a (7.41) se obtiene la solución de (7.36) con condiciones iniciales (7.39)

$$y(t) = L^{-1}\left\{\frac{F(s)}{P(s)}\right\} + L^{-1}\left\{\frac{P_0(s)}{P(s)}\right\} \tag{7.42}$$

Para dar la interpretación física a (7.41) y (7.42), consideremos dos problemas más sencillos. Con la notación $D = d/dt$, el primer problema es

$$P(D)\, u = f(t), \quad u(0) = u^{(1)}(0) = \cdots = u^{(n-1)}(0) = 0$$

que corresponde a un oscilador forzado con condiciones iniciales cero. El segundo problema es

$$P(D)\,v = 0, \quad v(0) = y_0,\ v^{(1)}(0) = y_1,\ \ldots,\ v^{(n-1)}(0) = y_{n-1}$$

que corresponde a un oscilador libre con condiciones iniciales arbitrarias. Según la fórmula (7.41), la solución del primer problema es

$$U(s) = \frac{F(s)}{P(s)}, \quad u(t) = L^{-1}\left\{\frac{F(s)}{P(s)}\right\}$$

y la del segundo es

$$V(s) = \frac{P_0(s)}{P(s)}, \quad v(t) = L^{-1}\left\{\frac{P_0(s)}{P(s)}\right\}$$

Las fórmulas generales (7.41) y (7.42) nos indican que

$$Y(s) = U(s) + V(s), \quad y(t) = u(t) + v(t) \tag{7.43}$$

es decir, $y(t)$ es la suma de dos funciones más sencillas, $u(t)$ y $v(t)$. Para sistemas disipativos, $u(t)$ representa un estado estable, mientras que $v(t)$ representa procesos transitorios surgidos por las condiciones iniciales. La separación de los efectos estables y transitorios es una de las muchas ventajas del método de la transformada de Laplace. Cabe mencionar que el paso final, la restauración de la solución $y(t)$ a partir de la transformada $Y(s)$, en muchas aplicaciones tecnológicas no es de importancia primordial. A menudo resulta que la parte principal del estudio de un sistema ocurre en el espacio de frecuencia, es decir, en el espacio de la variable $s$, pero no en el espacio de tiempo $t$.

**Ejemplo 7.11.** *Hallamos la solución de dos problemas con valores iniciales*

$$\ddot{y} + 2\beta\dot{y} + \omega^2 y = \left\{\begin{array}{l} f_1(t) \\ f_2(t) \end{array}\right., \quad y(0) = 1,\ \dot{y}(0) = 0$$

*en donde $\beta$ y $\omega$ son constantes, $f_1(t) = \cos\Omega t$ para $0 < t$, $f_2(t) = 1$ para $0 < t < a$ y es igual a cero en caso contrario. Dicha ecuación describe un oscilador amortiguado que es forzado, en primer caso, por la fuerza armónica $f_1(t)$ y, en segundo caso, por un impulso unitario $f_2(t)$; las condiciones iniciales son idénticas para ambos casos. Es fácil ver que $f_2(t) = u_0(t) - u_a(t)$, en donde $u_a(t)$ es la función escalón unidad. Mediante la transformación de Laplace (véanse las fórmulas (7.15), (7.23) y (7.28)), los problemas se reducen a las ecuaciones subsidiarias para $Y(s) = L\{y(t)\}$*

$$\left(s^2Y - s\right) + 2\beta(sY - 1) + \omega^2 Y = \left\{\begin{array}{l} s/\left(s^2 + \Omega^2\right) \\ 1/s - e^{-as}/s \end{array}\right.$$

*cuyas soluciones son*

$$Y_1(s) = \frac{F_1(s)}{P(s)} + \frac{P_0(s)}{P(s)}, \quad Y_2(s) = \frac{F_2(s)}{P(s)} + \frac{P_0(s)}{P(s)}$$

*en donde* $P(s) = s^2 + 2\beta s + \omega^2$ *es el polinomio característico,* $P_0(s) = s + 2\beta$ *es el polinomio que depende de las condiciones iniciales,* $F_1(s) = L\{f_1(t)\} = s/\left(s^2+\Omega^2\right)$ *y* $F_2(s) = L\{f_2(t)\} = 1/s - e^{-as}/s$. *Sean* $\lambda_{1,2} = -\beta \pm \sqrt{\beta^2 - \omega^2} = -\beta \pm i\tilde{\omega}$ $\left(\tilde{\omega} \equiv \sqrt{\omega^2 - \beta^2}\right)$ *raíces de la ecuación característica,* $P(s) = 0$*; de esta manera,* $P(s) = (s - \lambda_1)(s - \lambda_2)$. *En términos de fracciones parciales, la parte transitoria de la solución es idéntica para los dos sistemas*

$$\frac{P_0(s)}{P(s)} = \frac{s + 2\beta}{(s-\lambda_1)(s-\lambda_2)} = \frac{2\beta + \lambda_1}{(s-\lambda_1)(\lambda_1 - \lambda_2)} - \frac{2\beta + \lambda_2}{(s-\lambda_2)(\lambda_1 - \lambda_2)}$$

*mientras que las vibraciones forzadas son diferentes: para el primer sistema vibrante, se tiene*

$$\begin{aligned}\frac{F_1(s)}{P(s)} &= \frac{s}{\left(s^2+\Omega^2\right)(s-\lambda_1)(s-\lambda_2)} \\ &= \frac{A}{(s+i\Omega)} + \frac{B}{(s-i\Omega)} + \frac{C}{(s-\lambda_1)} + \frac{D}{(s-\lambda_2)}\end{aligned}$$

*en donde las constantes A, B, C y D son*

$$\begin{aligned}A &= \frac{1}{2(i\Omega+\lambda_1)(i\Omega+\lambda_2)}, \quad B = \frac{1}{2(i\Omega-\lambda_1)(i\Omega-\lambda_2)} \\ C &= \frac{\lambda_1}{\left(\lambda_1^2+\Omega^2\right)(\lambda_1-\lambda_2)}, \quad D = \frac{\lambda_2}{\left(\lambda_2^2+\Omega^2\right)(\lambda_2-\lambda_1)}\end{aligned}$$

*para el segundo oscilador, se tiene*

$$\begin{aligned}\frac{F_2(s)}{P(s)} &= \frac{1-e^{-as}}{s(s-\lambda_1)(s-\lambda_2)} \\ &= \left(1 - e^{-as}\right)\left[\frac{K_0}{s} + \frac{K_1}{s-\lambda_1} + \frac{K_2}{s-\lambda_2}\right]\end{aligned}$$

*en donde*

$$K_0 = (\lambda_1\lambda_2)^{-1}, \quad K_1 = \lambda_1^{-1}(\lambda_1-\lambda_2)^{-1}, \quad K_2 = \lambda_2^{-1}(\lambda_2-\lambda_1)^{-1}$$

*Por lo tanto, mediante (7.12) se obtiene la parte transitoria*

$$y_0(t) = L^{-1}\left\{\frac{P_0(s)}{P(s)}\right\} = \frac{2\beta+\lambda_1}{\lambda_1-\lambda_2}\exp(\lambda_1 t) - \frac{2\beta+\lambda_2}{\lambda_1-\lambda_2}\exp(\lambda_2 t)$$

*que decae exponencialmente,* $\exp(-\beta t)$. *Por otra parte, en el caso de una fuerza externa armónica, el sistema vibrante permanecerá oscilando con la frecuencia de la fuerza, como se ve de la solución*

$$L^{-1}\left\{\frac{F_1(s)}{P(s)}\right\} = A\exp(-i\Omega t) + B\exp(i\Omega t) + C\exp(\lambda_1 t) + D\exp(\lambda_2 t)$$

*en donde también se usó (7.12). Si la fuerza externa actúa sólo por un periodo de tiempo $T = a$, el comportamiento de un sistema vibrante es diferente, como se ve de la solución*

$$\begin{aligned} L^{-1}\left\{\frac{F_2(s)}{P(s)}\right\} &= K_0 + K_1 \exp(\lambda_1 t) + K_2 \exp(\lambda_2 t) - \\ & [K_0 + K_1 \exp(\lambda_1 (t-a)) + K_2 \exp(\lambda_2 (t-a))]\, u_a(t) \end{aligned}$$

*que, a fin de cuentas decae exponencialmente,* $\exp(-\beta t)$.

**Ejemplo 7.12.** *Hallamos la solución del sistema de ecuaciones*

$$\begin{aligned} \dot{y} - 2y + z &= 0 \\ \dot{z} - y - 2z &= 0 \end{aligned}$$

*con condiciones iniciales* $y(0) = 1$ *y* $z(0) = 0$. *Aplicando la transformada de Laplace a cada una de las ecuaciones, se obtienen las ecuaciones subsidiarias*

$$\begin{aligned} (s-2)\,Y(s) + Z(s) = 1 \\ Y(s) - (s-2)\,Z(s) = 0 \end{aligned}$$

*Despejando este sistema de ecuaciones algebraicas con respecto a* $Y(s)$ *y* $Z(s)$, *se tiene*

$$Y(s) = \frac{s-2}{(s-2)^2+1}, \quad Z(s) = \frac{1}{(s-2)^2+1}$$

*De donde, mediante (7.18) se obtiene la solución*

$$y(t) = L^{-1}\{Y(s)\} = e^{2t}\cos t, \quad z(t) = L^{-1}\{Z(s)\} = e^{2t}\operatorname{sen} t$$

### 7.4.2. Ecuaciones diferenciales ordinarias con coeficientes variables

Hemos visto, ecuación (7.31), que

$$L\{ty(t)\} = -Y'(s)$$

Aplicando esta fórmula $n$ veces consecutivamente a la transformada $L\{t^n f(t)\}$, es fácil comprobar que

$$L\{t^n y(t)\} = (-1)^n\, Y^{(n)}(s) \tag{7.44}$$

La última fórmula nos permite escribir la transformada de producto de $t^n$ por cualquier derivada de $f(t)$; por ejemplo,

$$L\left\{t^2 \frac{dy(t)}{dt}\right\} = \frac{d^2}{ds^2}[sY(s) - y(0)] = s\frac{d^2Y(s)}{ds^2} + 2\frac{dY(s)}{ds}$$

$$\begin{aligned} L\left\{t\frac{d^2y(t)}{dt^2}\right\} &= -\frac{d}{ds}\left[s^2Y(s) - sy(0) - y^{(1)}(0)\right] \\ &= -s^2\frac{dY(s)}{ds} - 2sY(s) + y(0) \end{aligned}$$

En consecuencia, una ecuación diferencial lineal con respecto a $y(t)$ cuyos coeficientes sean polinomios en $t$, se transforma en una ecuación diferencial con respecto a $Y(s)$ cuyos coeficientes sean polinomios en $s$. Cuando la ecuación transformada es más simple que la original, la resolución de la última nos permite encontrar la solución de la ecuación original.

Si los coeficientes de la ecuación original son polinomios de primer grado, la ecuación transformada es la ecuación diferencial lineal de primer orden cuya solución puede ser escrita en forma de una integral. Sin embargo, para hallar la solución de la ecuación original, hay que implementar la transformada inversa de la solución de la ecuación transformada.

**Ejemplo 7.13.** *Hallamos la solución del problema*

$$\ddot{y}(t) + t\dot{y}(t) - y(t) = 0, \quad y(0) = 0, \quad \dot{y}(0) = 1$$

*La ecuación transformada es*

$$s^2Y(s) - 1 - \frac{d}{ds}[sY(s)] - y(s) = 0$$

*o*

$$\frac{dY(s)}{ds} + \left(\frac{2}{s} - s\right)Y(s) = -\frac{1}{s}$$

*que es una ecuación diferencial lineal de primer orden. Multiplicando por el factor integrante*

$$\exp\left[\int\left(\frac{2}{s} - s\right)ds\right] = s^2\exp\left(-s^2/2\right)$$

*la ecuación transformada se escribe en la forma*

$$\frac{d}{ds}\left[s^2\exp\left(-s^2/2\right)Y(s)\right] = -s\exp\left(-s^2/2\right)$$

*La integración de esta última conduce a la solución general*

$$Y(s) = \frac{1}{s^2} + \frac{C}{s^2}\exp\left(s^2/2\right)$$

*en donde $C$ es una constante de integración. La función $Y(s)$ es la transformada de Laplace y, por tanto, debe desvanecerse cuando $s \to \infty$. Esto requiere que $C = 0$. Entonces, $Y(s) = 1/s^2$ y la solución del problema original es $y(t) = L^{-1}\left\{1/s^2\right\} = t$.*

### 7.4.3. Ecuaciones integrales del tipo de convolución

Una ecuación que comprende la función incógnita dentro de una integral se llama **ecuación integral**. Resulta que en varios problemas aplicados, la integral de dicha ecuación es la de convolución. Estas ecuaciones integrales se transforman en ecuaciones algebraicas mediante la transformada de Laplace.

La ecuación integral general del tipo de convolución tiene la forma

$$y(t) = f(t) + \int_0^t g(t-\tau)\, y(\tau)\, d\tau \tag{7.45}$$

en donde las funciones $f(t)$ y $g(t)$ son dadas, mientras que $y(t)$ es la función incógnita. La transformada de Laplace de la ecuación (7.45) conduce a la ecuación auxiliar

$$Y(s) = F(s) + G(s)\, Y(s)$$

de donde se obtiene la transformada de la función incógnita

$$Y(s) = \frac{F(s)}{1 - G(s)} \tag{7.46}$$

Por tanto, la solución de (7.45) es

$$y(t) = L^{-1}\left\{\frac{F(s)}{1 - G(s)}\right\} \tag{7.47}$$

Es fácil ver que dicho método es aplicable aun cuando la ecuación (7.45) es modificada por la sustitución de una combinación lineal de $y(t)$ y sus derivadas en lugar de $y(t)$. La transformada de la ecuación modificada es todavía una ecuación algebraica en $Y(s)$. Por ejemplo, la ecuación integral diferencial

$$Ay(t) + B\dot{y}(t) = f(t) + \int_0^t g(t-\tau)\, y(\tau)\, d\tau \tag{7.48}$$

en donde $A$ y $B$ son constantes, conduce a la ecuación transformada

$$(A + Bs)\, Y(s) - By(0) = F(s) + G(s)\, Y(s)$$

que es fácil despejar con respecto a $Y(s)$

$$Y(s) = \frac{F(s) + By(0)}{A + Bs - G(s)} \tag{7.49}$$

Otro ejemplo nos lo proporciona la ecuación

$$f(t) = \int_0^t (t-\tau)^{-b}\, \dot{y}(\tau)\, d\tau, \quad 0 < b < 1 \tag{7.50}$$

conocida como ecuación integral de Abel. La transformada de ésta es

$$F(s) = L\left\{t^{-b}\right\} L\{\dot{y}(t)\} = \frac{\Gamma(1-b)}{s^{1-b}}\, (sY(s) - y(0))$$

de donde se tiene

$$Y(s) = \frac{1}{\Gamma(1-b)} \frac{F(s)}{s^b} + \frac{y(0)}{s}$$

Por tanto, la solución de la ecuación integral de Abel es

$$\begin{aligned} y(t) &= y(0) + \frac{1}{\Gamma(b)\Gamma(1-b)} \left(t^{b-1} * f(t)\right) \\ &= y(0) + \frac{\operatorname{sen} \pi b}{\pi} \left(t^{b-1} * f(t)\right) \end{aligned} \quad (7.51)$$

en donde se usó la fórmula (7.16).

### 7.4.4. La integral de inversión de la transformada de Laplace

En las secciones anteriores, la transformada inversa de Laplace, definida como la correspondencia entre la transformada $F(s) = L\{f\}$ y la función $f(t)$, ha sido usada en la práctica en forma de tablas de correspondencia. Sin embargo, es de gran interés en las aplicaciones tener una fórmula que nos permita, sólo sabiendo una $F(s)$, reconstruir la función $f(t)$ correspondiente. Con base en cierta semejanza entre la transformada de Laplace y la de Fourier, es de esperar que la transformada inversa de Laplace también sea representada por una integral.

Con el fin de establecer la fórmula deseada, hacemos la continuación analítica de $F(s)$ sobre todo el plano complejo, es decir, suponemos que $s = z = x + iy$ es una variable compleja. Por definición de la transformada, se tiene

$$\begin{aligned} F(z) &= F(x+iy) = \int_0^\infty f(t) e^{-zt} dt = \int_0^\infty f(t) e^{-(x+iy)t} dt \\ &= \int_0^\infty e^{-xt} f(t) e^{-iyt} dt, \quad x = \operatorname{Re} z > \gamma \end{aligned} \quad (7.52)$$

Definida la función

$$\phi(t) = e^{-xt} f(t) u_0(t)$$

la relación (7.52) puede interpretarse como la transformada de Fourier de $\phi(t)$

$$F(x+iy) = \int_{-\infty}^{\infty} \phi(t) e^{-iyt} dt$$

Si la función $\phi(t)$ satisface condiciones suficientes para que exista la transformada de Fourier de ésta, podemos escribir la transformada inversa

$$\phi(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(x+iy) e^{iyt} dy$$

Usando la definición de $\phi(t)$, se tiene

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(x+iy) e^{(x+iy)t} dy$$

para $t > 0$. Dado que en las operaciones anteriores, $x$ está considerada como una constante, digamos $x = c$, se tiene que $dz = d(x + iy) = idy$ y

$$f(t) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} F(z)\, e^{zt} dz, \quad t > 0 \tag{7.53}$$

en donde la trayectoria de integración es la línea recta vertical $x = c > \gamma$ en el plano complejo $z$. La integral (7.53) representa la transformada inversa de Laplace en forma explícita. Se dice que una función $\chi(z)$ es de orden de $z^k$ y se escribe $O(z^k)$, si

$$\lim_{z\to\infty} \left|z^{-k} F(z)\right| < M$$

en donde $M$ es un número positivo. Si $F(z)$ es analítica de orden $O(z^{-k})$ en el semiplano $x > c$, donde $k > 1$, entonces se puede comprobar mediante el teorema integral de Cauchy que dicha integral no depende de $c > \gamma$, es decir, la trayectoria de integración es cualquier línea vertical con $c > \gamma$.

Para evaluar la integral (7.53), es conveniente aplicar el teorema de residuos. Sea $F(z)$ una función analítica de orden $O(z^{-k})$, donde $k > 1$ para toda $z$, excepto un conjunto de puntos singulares aislados $s_1, s_2, \ldots, s_N$ que se encuentran en el semiplano $x < \gamma$. Dado que $e^{zt}$ es una función analítica de $z$, los puntos singulares de $F(z)$ son los del integrando $F(z)\, e^{zt}$ en la integral de inversión. Según el teorema de residuos, la integral de $F(z)\, e^{zt}$ a lo largo de una trayectoria $C$ que encierre los puntos singulares $s_1, s_2, \ldots, s_N$ es igual a la suma de los residuos multiplicada por $2\pi i$

$$\oint_C F(z)\, e^{zt} dz = 2\pi i \sum_{n=1}^{N} \text{Res}\left(F(s_n)\, e^{s_n t}\right) = 2\pi i \sum_{n=1}^{N} e^{s_n t}\, \text{Res}\,(F(s_n))$$

Consideremos la trayectoria de integración $C$ compuesta de un segmento de la recta que une los puntos $c - iA$ y $c + iA$ y de un semicírculo $C_A$ que une dichos puntos, encerrando de esta manera los puntos singulares. Por tanto,

$$\int_{c-iA}^{c+iA} F(z)\, e^{zt} dz + \oint_{C_A} F(z)\, e^{zt} dz = 2\pi i \sum_{n=1}^{N} e^{s_n t}\, \text{Res}\,(F(s_n))$$

Cuando $A \to \infty$, bajo ciertas condiciones sobre la función $F(z)$, la integral a lo largo del contorno $C_A$ tiende a cero, quedando

$$\int_{c-i\infty}^{c+i\infty} F(z)\, e^{zt} dz = 2\pi i \sum_{n=1}^{N} e^{s_n t}\, \text{Res}\,(F(s_n))$$

o

$$f(t) = L^{-1}\{F(s)\} = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} F(z)\, e^{zt} dz = \sum_{n=1}^{N} e^{s_n t}\, \text{Res}\,(F(s_n)) \tag{7.54}$$

en donde $s_1, s_2, \ldots, s_N$ son todos los puntos singulares de $F(z)$.

En la práctica es conveniente usar la fórmula (7.54) de manera formal, sin respetar las condiciones que la integral a lo largo de $C_A$ tiende a cero, cuando $A \to \infty$. Al obtener de esta manera una función $f(t)$, tiene que averiguar si ésta es una solución del problema al resolver por el método de la transformada de Laplace.

**Ejemplo 7.14.** *Hallamos la transformada inversa de* $F(s) = e^{\gamma s}/\left(s^2+a^2\right)$. *La función*

$$F(z)\,e^{zt} = \frac{e^{z(\gamma+t)}}{s^2+a^2}$$

*tiene polos simples en* $s_1 = -ia$ *y* $s_1 = ia$. *Los residuos correspondientes son*

$$\frac{e^{-ia(\gamma+t)}}{2ia}, \quad \frac{e^{ia(\gamma+t)}}{2ia}$$

*Por tanto, la transformada inversa es*

$$f(t) = L^{-1}\{F(s)\} = \frac{1}{a}\left(\frac{e^{ia(\gamma+t)} - e^{-ia(\gamma+t)}}{2i}\right) = \frac{1}{a}\operatorname{sen} a(\gamma+t)$$

## 7.5. Fracciones parciales

En muchas aplicaciones importantes, como hemos visto en las secciones anteriores, la solución de una ecuación subsidiaria $Y(s)$ es una función racional, es decir, la razón de dos polinomios $F(s)$ y $P(s)$ en $s$

$$Y(s) = \frac{F(s)}{P(s)}$$

En tal caso, es posible determinar $y(t) = L^{-1}\{Y\}$, la solución del problema original, expresando $Y(s)$ en términos de fracciones parciales, dado que las transformadas inversas correspondientes son bien conocidas. En realidad, lo hemos hecho en varios ejemplos de este capítulo en una forma implícita. En esta sección presentamos la expansión en fracciones parciales de manera más sistemática, considerando los casos más frecuentes en las aplicaciones.

En un principio, suponemos que $F(s)$ y $P(s)$ son polinomios con coeficientes reales y no tienen factores comunes, el grado de $F(s)$ es menor que el de $P(s)$.

**Caso 1. Raíces distintas**

Si $P(s)$ es un polinomio de grado $n$ y $\lambda_1, \lambda_2, \ldots, \lambda_n$ son sus raíces reales o complejas distintas, entonces se tiene la factorización $P(s) = (s-\lambda_1)(s-\lambda_2)\cdots(s-\lambda_n)$. Si $F(s)$ es cualquier polinomio de grado menor que $n$, entonces existen las constantes $a_k$ $(k = 1, 2, \ldots, n)$ tales que es válida la expansión en fracciones parciales

$$\frac{F(s)}{P(s)} = \frac{F(s)}{(s-\lambda_1)(s-\lambda_2)\cdots(s-\lambda_n)} = \frac{a_1}{s-\lambda_1} + \frac{a_2}{s-\lambda_2} + \cdots + \frac{a_n}{s-\lambda_n} \quad (7.55)$$

para $s \neq \lambda_k$. Las constantes $a_k$ pueden determinarse multiplicando ambas partes de esta relación por $P(s) = (s-\lambda_1)(s-\lambda_2)\cdots(s-\lambda_n)$ e igualando los coeficientes de potencias de $s$. Sin embargo, el procedimiento más eficaz consiste en multiplicar por $(s-\lambda_k)$ y, después, tender $s \to \lambda_k$. La ilustración es como sigue, por medio de un ejemplo.

Considérese

$$\frac{F(s)}{(s-\lambda_1)(s-\lambda_2)(s-\lambda_3)} = \frac{a_1}{s-\lambda_1} + \frac{a_2}{s-\lambda_2} + \frac{a_3}{s-\lambda_3}$$

en donde $F(s)$ es un polinomio de grado no más que dos. Multiplicando por $(s-\lambda_1)$, se tiene

$$\frac{F(s)}{(s-\lambda_2)(s-\lambda_3)} = a_1 + \left(\frac{a_2}{s-\lambda_2} + \frac{a_3}{s-\lambda_3}\right)(s-\lambda_1)$$

Cuando $s \to \lambda_1$, los términos que contienen $a_2$ y $a_3$ se desvanecen, quedando como resultado

$$a_1 = \frac{F(\lambda_1)}{(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)}$$

Las fórmulas para $a_2$ y $a_3$ son similares

$$a_2 = \frac{F(\lambda_2)}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)}$$

$$a_3 = \frac{F(\lambda_3)}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)}$$

**Caso 2. Raíces múltiples**

En el caso anterior se supuso que no hay raíces múltiples de $P(s)$. Al contrario, sea $\lambda_1$ la raíz de multiplicidad $m$, es decir, $P(s) = (s-\lambda_1)^m (s-\lambda_2)\cdots(s-\lambda_n)$. Sea $Y(s) = F(s)/P(s)$ una función racional, en donde el polinomio $P(s) = (s-\lambda_1)^m (s-\lambda_2)\cdots(s-\lambda_n)$ es de grado $N = m+n-1$ y $F(s)$ es un polinomio de grado menor que $N$. Entonces, es válida la expansión

$$\frac{F(s)}{(s-\lambda_1)^m (s-\lambda_2)\cdots(s-\lambda_n)} \tag{7.56}$$

$$= \frac{a_{11}}{s-\lambda_1} + \frac{a_{12}}{(s-\lambda_1)^2} + \cdots + \frac{a_{1m}}{(s-\lambda_1)^m} + \frac{a_2}{s-\lambda_2} + \cdots + \frac{a_n}{s-\lambda_n}$$

para toda $s \neq \lambda_k$ $(k = 1, 2, \ldots, n)$. Los coeficientes son

$$a_{1l} = \frac{1}{(m-l)!} \frac{d^{m-l} Q_{\lambda_1}(s)}{ds^{m-l}}\bigg|_{s=\lambda_1}, \quad l = 1, 2, \ldots, m \tag{7.57}$$

en donde

$$Q_{\lambda_1}(s) = (s-\lambda_1)^m \frac{F(s)}{P(s)} = \frac{F(s)}{(s-\lambda_2)\cdots(s-\lambda_n)} \tag{7.58}$$

y

$$a_k = \frac{F(\lambda_k)}{(\lambda_k-\lambda_1)^m(\lambda_k-\lambda_2)\cdots(\lambda_k-\lambda_{k-1})(\lambda_k-\lambda_{k+1})\cdots(\lambda_k-\lambda_n)}$$
$$= (s-\lambda_k)\left.\frac{F(s)}{P(s)}\right|_{s=\lambda_k}, \quad k=2,3,\ldots,n \tag{7.59}$$

Para comprobar la relación (7.56), multiplicamos ambas partes de ésta por $P(s)$, quedando con la igualdad de dos polinomios de orden menor que $N$, $F(s)$ y un polinomio que contiene los $N$ coeficientes $a_{1l}$ y $a_k$ como incógnitas. Al igualar los coeficientes de dichos polinomios se obtienen $N$ ecuaciones con respecto a las $N$ incógnitas $a_{1l}$ y $a_k$. Despejando éstas se obtienen $a_{1l}$ y $a_k$, comprobando de esta manera la relación (7.56). Sin embargo, el procedimiento más eficaz para calcular los coeficientes es como sigue.

Para obtener $a_k$ se multiplica (7.56) por $(s-\lambda_k)$ y, tendiendo $s\to\lambda_k$, se obtiene (7.59).

Para obtener $a_{1l}$ se multiplica (7.56) por $(s-\lambda_1)^m$ y, después, se deriva $(m-l)$ veces, quedando

$$\frac{d^{m-l}Q_{\lambda_1}(s)}{ds^{m-l}} = \frac{(m-1)!}{(m-l)!}(s-\lambda_1)^{l-1}a_{11}+\frac{(m-2)!}{(m-l)!}(s-\lambda_1)^{l-2}a_{12}+\cdots+$$
$$(m-l)!a_{1l}+\frac{d^l}{ds^l}(s-\lambda_1)^m W \tag{7.60}$$

en donde

$$W(s)=\frac{a_2}{s-\lambda_2}+\cdots+\frac{a_n}{s-\lambda_n}$$

que es una función analítica en el punto $s=\lambda_1$. Entonces, para $s=\lambda_1$, de (7.60) se obtiene (7.57).

Nótese que el caso 1 se obtiene del caso 2 automáticamente, para $m=1$. Además, se conoce que

$$L^{-1}\left\{\frac{A}{(s-\lambda)^k}\right\}=A\frac{t^{k-1}}{(k-1)!}e^{\lambda t}, \quad k=1,2,\ldots \tag{7.61}$$

### Caso 3. Factores cuadráticos y raíces complejo conjugadas

A veces es conveniente usar factores cuadráticos en los denominadores de la expansión (7.55), en vez de factores lineales $(s-\lambda_k)$. La forma apropiada se obtiene mediante la agrupación del par de términos correspondientes en la expansión básica. Por ejemplo, el término con el factor cuadrático $(s-\lambda_1)(s-\lambda_2)$ en denominador, se obtiene de la siguiente manera:

$$\frac{a_1}{s-\lambda_1}+\frac{a_2}{s-\lambda_2} = \frac{(a_1+a_2)s-a_1\lambda_2-a_2\lambda_1}{(s-\lambda_1)(s-\lambda_2)}$$
$$\equiv \frac{As+B}{(s-\lambda_1)(s-\lambda_2)} \tag{7.62}$$

Las transformadas inversas correspondientes son

$$L^{-1}\left\{\frac{B}{(s-\lambda_1)(s-\lambda_2)}\right\} = \frac{B}{\lambda_1-\lambda_2}\left(e^{\lambda_1 t}-e^{\lambda_2 t}\right) \tag{7.63}$$

$$L^{-1}\left\{\frac{As}{(s-\lambda_1)(s-\lambda_2)}\right\} = \frac{A}{\lambda_1-\lambda_2}\left(\lambda_1 e^{\lambda_1 t}-\lambda_2 e^{\lambda_2 t}\right) \tag{7.64}$$

Del Álgebra se conoce que si el polinomio $P(s)$ tiene una raíz compleja $\lambda = \alpha + i\beta$, su conjugado $\lambda^* = \alpha - i\beta$ también es una raíz de $P(s)$. Entonces, la suma correspondiente de dos fracciones parciales en la expansión (7.55) se puede representar en la forma

$$\frac{a_1}{s-\lambda}+\frac{a_2}{s-\lambda^*}=\frac{As+B}{(s-\lambda)(s-\lambda^*)}$$

y, en consecuencia, la transformada inversa es

$$\begin{aligned} L^{-1}\left\{\frac{As+B}{(s-\lambda)(s-\lambda^*)}\right\} &= \frac{A}{i2\beta}\left(\lambda e^{\lambda t}-\lambda^* e^{\lambda^* t}\right)+\frac{B}{i2\beta}\left(e^{\lambda t}-e^{\lambda^* t}\right) \\ &= \frac{e^{\alpha t}}{\beta}\left[A(\alpha \operatorname{sen}\beta t+\beta\cos\beta t)+B\operatorname{sen}\beta t\right] \end{aligned}$$

**Ejemplo 7.15.** *Resuélvase el problema con valores iniciales*

$$y'''-y'=\operatorname{sen} t,\quad y(0)=2,\ y'(0)=0,\ y''(0)=1$$

*La ecuación transformada es*

$$s^3Y-2s^2-1-(sY-2)=\left(s^2+1\right)^{-1}$$

*Despejando $Y(s)$, se tiene*

$$Y(s)=\frac{2s^2-1}{s^3-s}+\frac{1}{(s^3-s)(s^2+1)}=\frac{2s^3+s}{(s^2-1)(s^2+1)}$$

*La expansión en fracciones parciales es*

$$\frac{2s^3+s}{(s^2-1)(s^2+1)}=\frac{a}{s-1}+\frac{b}{s+1}+\frac{c+ds}{s^2+1}$$

*Multiplicando por $(s-1)$ y tendiendo $s\to 1$, se tiene $a=3/4$. De manera similar, $b=3/4$. La manera sencilla de determinar los dos coeficientes restantes es asignar a $s$ dos valores particulares cualesquiera. Para $s=0$ se obtiene $c=0$. Por otra parte, multiplicando por $s$ y, después, tendiendo $s\to\infty$, se tiene $2=a+b+d$, pues $d=1/2$. Finalmente,*

$$Y(s)=\frac{3/4}{s-1}+\frac{3/4}{s+1}+\frac{s/2}{s^2+1}$$

*La transformada inversa nos da la solución del problema,*

$$y(t)=\frac{3}{4}e^t+\frac{3}{4}e^{-t}+\frac{1}{2}\cos t$$

## 7.6. Funciones periódicas

El hecho de que las funciones periódicas aparezcan en muchos problemas prácticos justifica el tema de la presente sección.

Sea $f(t)$ una función periódica de periodo $T$, es decir, $f(t+T) = f(t)$ para toda $t > 0$. Si $f(t)$ es seccionalmente continua sobre un intervalo de longitud $T$, su transformada de Laplace existe y puede escribirse como una serie de integrales sobre periodos sucesivos:

$$F(s) = L\{f(t)\} = \int_0^{\infty} f(t)\, e^{-st} dt = \sum_{k=0}^{\infty} \int_{kT}^{(k+1)T} f(t)\, e^{-st} dt$$

La sustitución $t = \tau + kT$, para cada intervalo $k$-ésimo de la integración $(k = 0, 1, 2, \ldots)$, conduce a la relación

$$\begin{aligned} F(s) &= \sum_{k=0}^{\infty} \int_{kT}^{(k+1)T} f(t)\, e^{-st} dt = \left( \sum_{k=0}^{\infty} e^{-ksT} \right) \int_0^T f(\tau)\, e^{-s\tau} d\tau \\ &= \frac{1}{1 - e^{-sT}} \int_0^T f(\tau)\, e^{-s\tau} d\tau, \qquad s > 0 \end{aligned} \tag{7.65}$$

en donde hemos usado la fórmula para la suma de una progresión geométrica.

**Ejemplo 7.16.** *Hallamos la transformada de la onda cuadrada*

$$f(t) = \begin{cases} A & si \quad kT < t < (k+1)T \\ -A & si \ (k+1)T < t < (k+2)T \end{cases}$$

*donde* $k = 0, 1, 2, \ldots$ *Dado que el periodo es* $2T$*, al utilizar (7.65) se obtiene*

$$\begin{aligned} F(s) &= \frac{1}{1 - e^{-s2T}} \left( \int_0^T A e^{-s\tau} d\tau - \int_T^{2T} A e^{-s\tau} d\tau \right) \\ &= \frac{A\left(1 - e^{-sT}\right)}{s\left(1 + e^{-sT}\right)} = \frac{A}{s} \tanh \frac{sT}{2} \end{aligned}$$

**Ejemplo 7.17.** *Hallamos la transformada de la onda diente de sierra*

$$g(t) = \frac{A}{T}(t - kT) \quad cuando \quad kT < t < (k+1)T$$

*en donde* $k = 0, 1, 2, \ldots$ *El periodo de la función es* $T$*. Al integrar por partes se tiene*

$$\int_0^T \tau e^{-s\tau} d\tau = -\frac{T}{s} e^{-sT} - \frac{1}{s^2}\left(e^{-sT} - 1\right)$$

*y, a partir de (7.65), se obtiene el resultado*

$$G(s) = \frac{A}{Ts^2} - \frac{A e^{-sT}}{s\left(1 - e^{-sT}\right)}$$

**Ejemplo 7.18.** *La función escalera se define por*

$$f(t) = kA \quad \textit{cuando} \quad kT < t < (k+1)T$$

*en donde* $k = 0, 1, 2, \ldots$ *El periodo de la función es* $T$. *Notando que la función escalera* $f(t)$ *es la diferencia entre* $h(t) = (A/T)\,t$, *cuya transformada es* $H(s) = \left(A/Ts^2\right)$, *y la onda diente de sierra,* $g(t) = (A/T)(t - kT)$, *se obtiene la transformada*

$$F(s) = H(s) - G(s) = \frac{Ae^{-sT}}{s\left(1 - e^{-sT}\right)}$$

**Nota bibliográfica:** para ampliar los conocimientos sobre el tema de este capítulo, puede consultar [1], [4], [6] y [11].

# Capítulo 8

# Ecuaciones diferenciales parciales

Una ecuación que comprende una o más derivadas parciales de una o más variables independientes se llama **ecuación diferencial parcial** (EDP). Las ecuaciones diferenciales parciales aparecen en diversos problemas físicos y geométricos, cuando las funciones que intervienen dependen de dos o más variables independientes. Estas variables pueden ser el tiempo y una o más coordenadas en el espacio, o todas pueden ser coordenadas en el espacio. El **orden de una ecuación diferencial parcial** es el orden de la derivada mayor. Puesto que $u = u(x)$ es una función de $n$ variables independientes, $x = (x_1, \ldots, x_n)$, la relación funcional

$$F\left(x, u, grad\, u, \frac{\partial^2 u}{\partial x_i \partial x_j}\right) = 0 \tag{8.1}$$

es una ecuación diferencial parcial de segundo orden. Se dice que la EDP es **lineal** si es de primer grado en la variable dependiente (la función desconocida) y sus derivadas parciales. Si cada uno de los términos de una ecuación de este tipo contiene a la función desconocida o una de sus derivadas, se dice que es **homogénea**; de lo contrario es **no homogénea**.

Una **solución** de una ecuación diferencial parcial en alguna región $R$ del espacio de variables independientes es una función que tiene todas las derivadas parciales que aparecen en la ecuación, en algún dominio que contenga a $R$, y que satisface a tal ecuación en todo punto de $R$. En general, la totalidad de las soluciones de una ecuación diferencial parcial es muy grande. Por ejemplo, las funciones

$$u = x^2 - y^2, \quad u = e^x \cos y, \quad u = \ln\left(x^2 + y^2\right)$$

que son completamente diferentes entre sí, son soluciones de la ecuación de Laplace

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

Más adelante se ve que la solución única de la EDP correspondiente a un problema físico dado se obtiene mediante la aplicación de información adicional proveniente de la situación física. En algunos casos se dan los valores de la solución sobre la frontera de algún dominio (**condiciones en la frontera**); en otros casos, cuando el tiempo $t$ es una de las variables independientes, se indican los valores de la solución en un instante $t = t_0$ (**condiciones iniciales**).

Las soluciones de las EDP lineales homogéneas poseen una propiedad idéntica a las de las EDO lineales homogéneas. Si $u_1$ y $u_2$ son soluciones cualesquiera de una EDP lineal homogénea, entonces $u = Au_1 + Bu_2$ también es una solución de esa ecuación, donde $A$ y $B$ son constantes. La comprobación de esta propiedad es muy sencilla.

## 8.1. Clasificación de EDP lineales de segundo orden

En la Física, una gran variedad de problemas se reduce a las EDP lineales de segundo orden

$$\sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_i(x) \frac{\partial u}{\partial x_i} + c(x) u = F(x) \tag{8.2}$$

donde $u = u(x)$ es la función desconocida de variables independientes $x = (x_1, \ldots, x_n)$. Los coeficientes $a_{ij}(x)$, $b_i(x)$, $c(x)$ y el término libre $F(x)$ son funciones continuas de $x = (x_1, \ldots, x_n)$.

Antes de empezar con el planteamiento y la solución de diferentes problemas físicos, es necesario clasificar las EDP correspondientes. Con este fin reescribimos la ecuación (8.2) en forma más compacta

$$\sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + \Phi(x, u, \operatorname{grad} u) = 0 \tag{8.3}$$

donde

$$\Phi(x, u, \operatorname{grad} u) \equiv \sum_{i=1}^{n} b_i(x) \frac{\partial u}{\partial x_i} + c(x) u - F(x)$$

Averiguamos cómo se cambiarán los coeficientes $a_{ij}(x)$ a un cambio no singular de variables independientes $y = y(x)$, es decir,

$$y_i = y_i(x_1, \ldots, x_n), \quad J\left(\frac{y_1, y_2, \ldots, y_n}{x_1, x_2, \ldots, x_n}\right) \neq 0, \quad i = 1, 2, \ldots, n \tag{8.4}$$

Dado que el Jacobiano de esta transformación es $J \neq 0$, las variables $x = (x_1, \ldots, x_n)$ se pueden expresar a través de las variables $y = (y_1, \ldots, y_n)$, $x = x(y)$, en una vecindad de cada punto $x$. Sea $u(x(y)) = \tilde{u}(y)$, es decir, $u(x) = \tilde{u}(y(x))$. Entonces,

$$\frac{\partial u}{\partial x_i} = \sum_{l=1}^{n} \frac{\partial \tilde{u}}{\partial y_l} \frac{\partial y_l}{\partial x_i} \tag{8.5}$$

$$\frac{\partial^2 u}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_j}\left(\frac{\partial u}{\partial x_i}\right)$$
$$= \sum_{l,k=1}^{n} \frac{\partial^2 \tilde{u}}{\partial y_k \partial y_l} \frac{\partial y_l}{\partial x_i} \frac{\partial y_k}{\partial x_j} + \sum_{l=1}^{n} \frac{\partial \tilde{u}}{\partial y_l} \frac{\partial^2 y_l}{\partial x_i \partial x_j} \tag{8.6}$$

Sustituyendo las expresiones (8.5) y (8.6) en la ecuación (8.3), se tiene

$$\sum_{l,k=1}^{n} \frac{\partial^2 \tilde{u}}{\partial y_k \partial y_l} \sum_{i,j=1}^{n} a_{ij} \frac{\partial y_l}{\partial x_i} \frac{\partial y_k}{\partial x_j} + \sum_{l=1}^{n} \frac{\partial \tilde{u}}{\partial y_l} \sum_{i,j=1}^{n} a_{ij} \frac{\partial^2 y_l}{\partial x_i \partial x_j} + \Phi\left(y, \tilde{u}, \operatorname{grad} \tilde{u}\right) = 0 \tag{8.7}$$

Redefiniendo

$$\tilde{a}_{lk}\left(y\right) = \sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial y_l}{\partial x_i} \frac{\partial y_k}{\partial x_j} \tag{8.8}$$

$$\tilde{\Phi}\left(y, \tilde{u}, \operatorname{grad} \tilde{u}\right) \equiv \check{\Phi}\left(y, \tilde{u}, \operatorname{grad} \tilde{u}\right) + \sum_{l=1}^{n} \frac{\partial \tilde{u}}{\partial y_l} \sum_{i,j=1}^{n} a_{ij}\left(x\right) \frac{\partial^2 y_l}{\partial x_i \partial x_j} \tag{8.9}$$

de la ecuación (8.7) se obtiene la ecuación

$$\sum_{l,k=1}^{n} \tilde{a}_{lk}\left(y\right) \frac{\partial^2 \tilde{u}}{\partial y_k \partial y_l} + \tilde{\Phi}\left(y, \tilde{u}, \operatorname{grad} \tilde{u}\right) = 0 \tag{8.10}$$

que es también la ecuación diferencial parcial lineal de segundo orden en variables independientes $y = (y_1, \ldots, y_n)$.

En un punto dado $x_0$ se tiene $y_0 = y\left(x_0\right)$. Haciendo la notación

$$b_{li} = \frac{\partial y_l\left(x_0\right)}{\partial x_i}$$

de la ecuación (8.8) se obtiene la relación

$$\tilde{a}_{lk}\left(y_0\right) = \sum_{i,j=1}^{n} a_{ij}\left(x_0\right) \frac{\partial y_l}{\partial x_i} \frac{\partial y_k}{\partial x_j} = \sum_{i,j=1}^{n} b_{li} a_{ij}\left(x_0\right) b_{kj} \tag{8.11}$$

que en forma matricial se escribe como

$$\tilde{A} = BAB^T \tag{8.12}$$

donde $\tilde{A} = \left[\tilde{a}_{lk}\left(y_0\right)\right]$ y $A = \left[a_{ij}\left(x_0\right)\right]$ son matrices de coeficientes de segundas derivadas parciales, $B = [b_{li}]$ es la matriz de primeras derivadas parciales $\partial y_l / \partial x_i$ en el punto $x_0$ y su determinante es el Jacobiano de la transformación de coordenadas $y = y\left(x\right)$, $J\left(y/x\right) = \det B \neq 0$. Se conoce del Álgebra Lineal que a una matriz simétrica real $A$ se le puede diagonalizar mediante una transformación del tipo (8.12), en donde $B^T$ es una matriz ortogonal cuyas columnas son eigenvectores de $A$. La matriz diagonalizada $\tilde{A} = diag A$ tendrá en su diagonal principal los eigenvalores de $A$ repetidos de acuerdo con sus multiplicidades

algebraicas. Entonces, escogiendo la transformación lineal no singular de variables independientes $y = y(x)$ de tal manera que la matriz $B$ satisfaga dichas condiciones en el punto $x_0$, podemos simplificar la ecuación (8.3) en el punto $x_0$ reduciéndola a la forma

$$\sum_{l=1}^{r} \lambda_l \frac{\partial^2 \tilde{u}}{\partial^2 y_l} - \sum_{l=r+1}^{m} \lambda_l \frac{\partial^2 \tilde{u}}{\partial^2 y_l} + \tilde{\Phi}(y, \tilde{u}, \operatorname{grad} \tilde{u}) = 0, \qquad m \leq n \tag{8.13}$$

donde separamos los términos con eigenvalores positivos de aquellos con eigenvalores negativos de $A$. Introduciendo nuevas variables independientes $\tilde{y}_l = y_l/\sqrt{\lambda_l}$, finalmente llegamos a la **forma canónica** de una EDP lineal de segundo orden

$$\sum_{l=1}^{r} \frac{\partial^2 \tilde{u}}{\partial^2 \tilde{y}_l} - \sum_{l=r+1}^{m} \frac{\partial^2 \tilde{u}}{\partial^2 \tilde{y}_l} + \tilde{\tilde{\Phi}}(\tilde{y}, \tilde{u}, \operatorname{grad} \tilde{u}) = 0, \qquad m \leq n \tag{8.14}$$

Los números enteros $r$ y $m$ no dependen de la transformación $y = y(x)$.

La clasificación de las EDP lineales de segundo orden está basada en la forma canónica de una EDP en cada punto $x_0$ y, por tanto, es local. En otras palabras, las EDP se pueden clasificar con base en la matriz de coeficientes $A = [a_{ij}(x_0)]$.

Cuando en la ecuación (8.14) $m = n$ y todos los términos con segundas derivadas parciales son del mismo signo, es decir, $r = m$ o $r = 0$, la matriz $A = [a_{ij}(x_0)]$ está definida (positiva o negativa) y la EDP (8.3) es de **tipo elíptico**.

La ecuación diferencial parcial lineal de secundo orden, ecuación (8.3), es de **tipo hiperbólico** cuando $m = n$, $1 \leq r \leq n-1$ (es hiperbólica normal, si $m = n$, $r = n-1$).

La ecuación (8.3) es de **tipo parabólico** cuando $m < n$ (es parabólica normal, si $r = n-1$).

**Ejemplo 8.1.** *Sea $u = u(x_1, x_2, x_3)$ una función desconocida de tres variables independientes $x_1, x_2, x_3$ que satisface la EDP(2) lineal*

$$\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial x_3^2} + \Phi(x, u, gradu) = 0$$

*Entonces, dicha ecuación es del tipo elíptico en todo el espacio de las variables $x_1, x_2, x_3$. Cambiando sólo el signo de un término, de la ecuación anterior se obtiene la ecuación*

$$\frac{\partial^2 u}{\partial x_1^2} = \frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial x_3^2} + \Phi(x, u, gradu)$$

*que ya es una ecuación hiperbólica normal. Además, la ecuación*

$$\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \Phi(x, u, gradu) = 0$$

*es parabólica normal, dado que $r = n-1$.*

Hay que hacer notar que dicha clasificación depende del punto $x_0$, dado que los números $r$ y $m$ dependen de este punto. Por ejemplo, la **ecuación de Tricomi**

$$y\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \tag{8.15}$$

donde $u = u(x, y)$, es de tipo mixto, dependiendo del valor de $y$. Será hiperbólica si $y < 0$; elíptica si $y > 0$; o parabólica si $y = 0$.

## 8.2. Ecuaciones básicas de la Física Matemática

En muchas ocasiones la descripción matemática de procesos físicos o biológicos se reduce a las ecuaciones diferenciales o integrales lineales, o ecuaciones íntegro-diferenciales. Una clase amplia de problemas físicos se reduce a ecuaciones diferenciales parciales lineales de segundo orden, ecuación (8.2). En esta sección se describen ecuaciones típicas de la Física Matemática, omitiendo la deducción de éstas.

### 8.2.1. Ecuación de onda

Muchos problemas de mecánica (vibraciones de cuerdas, barras, membranas y volúmenes) y electromagnetismo (ondas electromagnéticas) llevan a la **ecuación de onda** de la forma

$$\rho\frac{\partial^2 u}{\partial t^2} = \operatorname{div}(p \operatorname{grad} u) - Ru_t - qu + F(x, t) \tag{8.16}$$

donde la función desconocida $u(x, t)$ depende de $n$ coordenadas espaciales $x = (x_1, \ldots, x_n)$ y del tiempo $t$. Los coeficientes $\rho(x) > 0$, $p(x) > 0$, $q(x)$ y $R$ son determinados por las propiedades del medio de propagación de la onda y, en general, dependen de $x$. El término libre $F(x, t)$ describe la intensidad de una fuerza externa. Según la definición de los operadores div y grad, se tiene

$$\operatorname{div}(p \operatorname{grad} u) = \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\left(p\frac{\partial u}{\partial x_i}\right)$$

La ecuación (8.16) es del tipo hiperbólico. Algunos de los casos particulares de esta ecuación son como siguen.

Las pequeñas vibraciones transversales de una cuerda elástica sin fricción se rigen por la **ecuación unidimensional de onda**

$$\rho\frac{\partial^2 u}{\partial t^2} = T_0\frac{\partial^2 u}{\partial x^2} + F(x, t) \tag{8.17}$$

donde la función $u(x, t)$ de una sola variable $x$ y del tiempo $t$, describe el desplazamiento transversal de cada punto de la cuerda, la tensión en la cuerda $T_0$ se supone constante, $\rho(x)$ es la densidad lineal de masa y $F(x, t)$ es una fuerza

externa. Si la densidad de masa es constante, $\rho(x) = \rho = const$, entonces, la ecuación de las vibraciones transversales de una cuerda elástica toma la forma

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2} + f(x,t) \tag{8.18}$$

donde $a = \sqrt{T_0/\rho}$ es la velocidad de onda transversal y $f(x,t) = F(x,t)/\rho$. Para obtener la solución única de la ecuación (8.17) hay que satisfacer las condiciones impuestas por el sistema físico. Es necesario conocer el desplazamiento $u(x,t)$ y la velocidad $\partial u(x,t)/\partial t$ en un momento inicial (condiciones iniciales) y, también, las condiciones en los extremos de la cuerda (condiciones en la frontera). Cuando los extremos de la cuerda son fijos, las condiciones en la frontera son $u(0,t) = u(L,t) = 0$.

De manera similar se obtiene la ecuación de pequeñas oscilaciones transversales de una membrana

$$\rho \frac{\partial^2 u}{\partial t^2} = T_0 \left( \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} \right) + F(x,t) \tag{8.19}$$

donde la función $u(x,t)$ de dos variables espaciales $x = (x_1, x_2)$ y del tiempo $t$ describe el desplazamiento transversal de cada punto de la membrana, la tensión por unidad de longitud $T_0$ se supone constante, $\rho(x)$ es la masa de la membrana por unidad de área y $F(x,t)$ es una fuerza externa. La ecuación (8.19) recibe el nombre de **ecuación bidimensional de onda**.

La **ecuación tridimensional de onda**

$$\frac{\partial^2 u}{\partial t^2} = a^2 \left( \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial x_3^2} \right) + F(x,t) \tag{8.20}$$

donde $x = (x_1, x_2, x_3)$, describe la propagación de ondas sonoras en un medio homogéneo e isotrópico y, también, de las ondas electromagnéticas en un medio homogéneo e isotrópico no conductor.

### 8.2.2. Ecuación de difusión y de Schrodinger

El flujo de calor y la difusión de partículas en un medio isotrópico se describe por la **ecuación de difusión general**

$$\rho \frac{\partial u}{\partial t} = \text{div}\,(p\,\text{grad}\, u) - qu + F(x,t) \tag{8.21}$$

que es una ecuación del tipo parabólico. Debido a que esta ecuación comprende a la primera derivada en tiempo $u_t$, en tanto que la ecuación de onda (8.16) contiene la segunda derivada $u_{tt}$, las soluciones de dichas ecuaciones serán por completo diferentes.

En el caso de conductividad térmica, la función desconocida $u(x,t)$ describe la distribución de la temperatura en el medio, $p(x)$ es la conductividad térmica, $c(x)$ es el calor específico, $\rho = c(x)\tilde{\rho}(x)$ y $\tilde{\rho}(x)$ es la densidad del material del medio, $q = 0$. La función $F(x,t)$ describe la intensidad de fuentes de calor en

el punto $x$ en el instante $t$. Si el medio es isotrópico y homogéneo, $\rho$ y $p$ son constantes y, definiendo el laplaciano $\nabla^2 = \Delta$, de la ecuación (8.21) se obtiene la **ecuación de conductividad térmica**

$$\frac{\partial u}{\partial t} = a^2 \Delta u + f(x,t) \tag{8.22}$$

donde $a^2 = p/\rho$, $f(x,t) = F(x,t)/\rho$.

La condición inicial para la ecuación de conductividad térmica (de difusión) comprende a la distribución inicial de temperatura (de concentración de partículas), $u_0(x) = u(x,0)$. En cuanto a las condiciones en la frontera $S$, se aplican diferentes relaciones de acuerdo con las condiciones físicas. Los ejemplos son como siguen.

a) Si en la frontera $S$ se mantiene una distribución de temperatura dada $u_1(x)$, entonces

$$u(x,t)\,|_S = u_1(x) \tag{8.23}$$

b) Si en $S$ se mantiene un flujo de calor dado $u_2(x)$, entonces

$$-p\,(\mathbf{n}\cdot\nabla u)\,|_S = u_2(x) \tag{8.24}$$

donde $\mathbf{n}$ es la normal exterior a la superficie $S$.

c) Se tienen también las condiciones en la frontera que son una combinación de las dos anteriores

$$p\,(\mathbf{n}\cdot\nabla u) + k\,[u(x,t) - u_3(x)]\,|_S = 0 \tag{8.25}$$

donde $k$ es el coeficiente de intercambio de calor en la frontera $S$ y $u_3(x)$ es la temperatura de ambiente en que está sumergido el sistema.

En el caso de la difusión de partículas en un medio, en la ecuación (8.21) la función desconocida $u(x,t)$ es la densidad (concentración) de partículas, $p(x)$ es el coeficiente de difusión, $\rho(x)$ es el coeficiente de porosidad del material del medio, $q$ describe la absorción de partículas por el medio. La función $F(x,t)$ describe la intensidad de fuentes de partículas en el punto $x$ en el instante $t$.

Nótese que la **ecuación de Schrodinger**

$$i\hbar\frac{\partial \psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\Delta\psi + V(x)\,\psi \tag{8.26}$$

que describe el movimiento de una partícula cuántica de masa $m$ en un campo potencial $V(x)$, es una ecuación diferencial parcial lineal de segundo orden del tipo parabólico, es decir, pertenece a la clase de la ecuación de difusión (8.21), en donde la variable de tiempo se sustituye por una variable imaginaria pura $it$. Aquí, $\psi(x,t)$ es la función de onda de dicha partícula cuántica y $\hbar$ es la constante de Planck.

### 8.2.3. Ecuaciones de Poisson, Laplace y Helmholtz

Una solución estacionaria es $u(x)$ que no depende del tiempo. Suponiendo que en las ecuaciones (8.16) y (8.21) las fuentes son estacionarias, $F(x,t) = F(x)$,

encontramos que las soluciones estacionarias $u(x)$ tanto de la ecuación de onda como de la de difusión satisfacen la ecuación estacionaria

$$\operatorname{div}(p \operatorname{grad} u) - qu + F(x) = 0 \tag{8.27}$$

que es del tipo elíptico.

Si $p$ es constante y $q = 0$, esta ecuación resulta en

$$\Delta u = -f(x) \tag{8.28}$$

que es la **ecuación de Poisson**, donde $f(x) = F(x)/p$. Si, además, $f(x) = 0$, se tiene la **ecuación de Laplace**

$$\Delta u = 0 \tag{8.29}$$

Considerando la ecuación de onda

$$\frac{\partial^2 u}{\partial t^2} = a^2 \Delta u + f(x,t)$$

suponemos que la fuerza externa es armónica

$$f(x,t) = f_0(x)\, e^{i\omega t}$$

de frecuencia $\omega$ y de amplitud $f_0(x)$. Si se busca la solución armónica de la misma frecuencia $\omega$ y de una amplitud desconocida $u_0(x)$

$$u(x,t) = u_0(x)\, e^{i\omega t}$$

de la ecuación de onda se obtiene

$$\Delta u_0 + k^2 u_0 = -\tilde{f}_0(x) \tag{8.30}$$

donde $k^2 = \omega^2/a^2$ y $\tilde{f}_0(x) = f_0(x)/a^2$. Esta última se conoce como **ecuación de Helmholtz**.

## 8.3. Problemas con condiciones en la frontera

Sea $D \subset R^n$ un dominio abierto en el espacio $R^n$ y sea $S$ su frontera seccionalmente continua, siendo $\bar{D} = D \cup S$. Para resolver un problema físico (una EDP correspondiente) que describe un proceso en la región $D$ del espacio, tiene que determinarse una solución $u(x,t)$ de la ecuación correspondiente. Dicha solución debe describir el estado inicial del proceso, es decir, satisfacer las condiciones iniciales y, además, satisfacer las condiciones en la frontera. La razón física para satisfacer estas condiciones es que, en realidad, se conoce o se prescribe el estado inicial del proceso (del sistema físico), y las condiciones en la frontera son las que describen el modo de interacción del sistema con el ambiente o el método de control del sistema. La razón matemática proviene de la necesidad de tener una solución única del problema. De hecho, la solución general de una

ecuación diferencial ordinaria de orden $n$ ya contiene $n$ constantes arbitrarias. Una solución de la EDP, por lo general, depende de una función arbitraria. Por ejemplo, la solución general de la ecuación

$$\frac{\partial^2 u(x,y)}{\partial x \partial y} = 0$$

es una función de la forma $u(x,y) = f(x) + g(y)$, en donde $f(x)$ y $g(y)$ son funciones arbitrarias. Debido a esto es necesario tener algunas condiciones adicionales, que son las condiciones iniciales y de la frontera, para discriminar una solución particular que describe un proceso real. La tarea de resolver una EDP con condiciones iniciales y de la frontera recibe el nombre de **problema de valores en la frontera** (**problema con valores frontera**). Se distinguen tres tipos de problemas de valores en la frontera.

1. El **problema de Cauchy** para las ecuaciones del tipo hiperbólico y parabólico se plantea de la siguiente manera. Se establecen condiciones iniciales; el dominio $D$ del proceso coincide con todo el espacio $R^n$ y no existen condiciones en la frontera (por no existir la frontera).

2. El problema de valores en la frontera para ecuaciones del tipo elíptico: se establecen condiciones en la frontera $S$ y no existen condiciones iniciales (puesto que el tiempo $t$ no está presente en la ecuación).

3. El **problema mixto** para las ecuaciones del tipo hiperbólico y parabólico: se establecen tanto condiciones iniciales como condiciones en la frontera y, además, $D \neq R^n$.

A continuación se consideran más a detalle estos planteamientos para las ecuaciones básicas de la Física Matemática.

## 8.3.1. Problema de Cauchy

Para las ecuaciones del tipo hiperbólico, en particular para la ecuación de onda

$$\rho \frac{\partial^2 u}{\partial t^2} = \operatorname{div}\left(p \operatorname{grad} u\right) - R u_t - q u + F(x,t) \tag{8.31}$$

el problema de Cauchy se plantea de la siguiente manera: hay que encontrar la función $u(x,t)$, $x = (x_1, x_2, \ldots, x_n)$ y $t \geqslant 0$, de la clase $C^2(t > 0) \cap C^1(t \geqslant 0)$, que satisface la ecuación (8.31) para $t > 0$ y cumple las condiciones iniciales para $t = 0$

$$u(x,t)\big|_{t=+0} = u_0(x), \qquad \left.\frac{\partial u(x,t)}{\partial t}\right|_{t=+0} = u_1(x) \tag{8.32}$$

donde $u_0(x)$ y $u_1(x)$ son funciones conocidas en todo el espacio. Además, es necesario que

$$F(x,t) \in C(t > 0), \quad u_0(x) \in C^1(R^n), \quad u_1(x) \in C(R^n)$$

Para las ecuaciones del tipo parabólico, en particular para la ecuación de difusión

$$\rho \frac{\partial u}{\partial t} = \operatorname{div}\left(p \operatorname{grad} u\right) - q u + F(x,t) \tag{8.33}$$

el problema de Cauchy es como sigue: hay que encontrar la función $u(x,t)$, $x = (x_1, x_2, \ldots, x_n)$ y $t \geqslant 0$, de la clase $C^1 (t > 0) \cap C(t \geqslant 0)$, que satisface la ecuación (8.33) para $t > 0$ y cumple la condición inicial para $t = 0$

$$u(x,t)|_{t=+0} = u_0(x) \tag{8.34}$$

Además, es necesario que $F(x,t) \in C(t > 0)$, $u_0(x) \in C(R^n)$.

### 8.3.2. Problemas con valores frontera para ecuaciones elípticas (problemas de Dirichlet y de Neumann)

En cuanto a las ecuaciones del tipo elíptico

$$\operatorname{div}(p \operatorname{grad} u) - qu + F(x) = 0 \tag{8.35}$$

sus soluciones no dependen de $t$ y el problema con valores en la frontera se plantea de la siguiente manera: encontrar la función $u(x)$, $x = (x_1, x_2, \ldots, x_n)$, de la clase $C^2(D) \cap C^1(\bar{D})$ que satisface la ecuación (8.35) en el dominio $D$ y, en la frontera $S$ satisface la condición de la forma

$$\left( \alpha u(x) + \beta \frac{\partial u(x)}{\partial \mathbf{n}} \right)\bigg|_{x \in S} = v(x) \tag{8.36}$$

donde $\partial u / \partial \mathbf{n} = (\mathbf{n} \cdot \nabla u)$ es la derivada normal de la solución, siendo $\alpha(x)$, $\beta(x)$ y $v(x)$ funciones dadas continuas en $S$, $\alpha \geq 0, \beta \geq 0, \alpha + \beta > 0$.

Se distinguen tres tipos de condiciones en la frontera, ecuación (8.36):

1. La condición en la frontera del primer tipo ($\alpha = 1$, $\beta = 0$):

$$u(x)|_{x \in S} = u_0(x) \tag{8.37}$$

2. La condición en la frontera del segundo tipo ($\alpha = 0$, $\beta = 1$):

$$\left.\frac{\partial u}{\partial \mathbf{n}}\right|_{x \in S} = u_1(x) \tag{8.38}$$

3. Condiciones en la frontera del tercer tipo ($\beta = 1$, $\alpha \geqslant 0$):

$$\left( \frac{\partial u}{\partial \mathbf{n}} + \alpha u \right)\bigg|_{x \in S} = u_2(x) \tag{8.39}$$

Para las ecuaciones de Laplace y Poisson, al problema con condiciones en la frontera del primer tipo

$$\Delta u = -f(x), \quad u(x)|_{x \in S} = u_0(x)$$

se le llama **problema de Dirichlet**. Mientras que para las ecuaciones de Laplace y Poisson, el problema con condición en la frontera del segundo tipo, $\partial u / \partial \mathbf{n}|_{x \in S} = u_1(x)$, se conoce como el **problema de Neumann**.

Sea $D$ un dominio abierto finito tal que el complemento $D_1 = R^n \backslash \bar{D}$ es un dominio infinito, donde $\bar{D} = D \cup S$. Se distinguen entre el problema de Dirichlet (de Neumann) interno y externo.

Para el problema de Dirichlet interno se busca la solución en un dominio finito $D$ que toma valores dados $u_0(x)$ en $S$. Para el problema de Dirichlet externo se busca la solución en un dominio infinito $D_1 = R^n \backslash \bar{D}$ que toma valores dados $u_0(x)$ en $S$ y tiende a cero en el infinito.

Para el problema de Neumann interno se busca la solución en un dominio finito $D$ que tiene una derivada normal dada $\partial u/\partial \mathbf{n}|_{x \in S} = u_1(x)$ en $S$. Para el problema de Neumann externo se busca la solución en un dominio infinito $D_1 = R^n \backslash \bar{D}$ que tiene una derivada normal dada $\partial u/\partial \mathbf{n}|_{x \in S} = u_1(x)$ en $S$ y tiende a cero en el infinito.

### 8.3.3. Problema mixto para ecuaciones hiperbólicas y parabólicas

Las soluciones de una ecuación hiperbólica o parabólica están definidas en un cilindro $C_\infty = D \otimes (t > 0)$ en el espacio $R^{n+1}$. El problema mixto para una ecuación hiperbólica, en particular para la ecuación (8.31), se plantea de la siguiente manera: hay que encontrar la función $u(x,t)$, $x = (x_1, x_2, \ldots, x_n) \in D$ y $t \geqslant 0$, de la clase $C^2(C_\infty) \cap C^1(\bar{C}_\infty)$, que satisface la ecuación (8.31) en el cilindro $C_\infty$, y cumple las condiciones iniciales para $t = 0$

$$u(x,t)|_{t=+0} = u_0(x), \quad \left.\frac{\partial u(x,t)}{\partial t}\right|_{t=+0} = u_1(x), \quad x \in D \tag{8.40}$$

y una condición en la frontera lateral del cilindro $C_\infty$ (en la frontera $S$ del dominio $D$ en cada instante $t$) de la forma

$$\left.\left(\alpha u(x,t) + \beta \frac{\partial u(x,t)}{\partial \mathbf{n}}\right)\right|_{x \in S} = v(x,t), \quad x \in S \tag{8.41}$$

donde $\partial u/\partial \mathbf{n} = (\mathbf{n} \cdot \nabla u)$ es la derivada normal de la solución, siendo $\alpha(x)$, $\beta(x)$ y $v(x)$ funciones dadas continuas en $S$, $\alpha \geq 0$, $\beta \geq 0$, $\alpha + \beta > 0$. También es necesario cumplir condiciones de suavidad de funciones

$$\begin{aligned} F(x,t) &\in C(C_\infty), \quad u_0(x) \in C^1(\bar{D}), \\ u_1(x) &\in C(\bar{D}), \quad v(x,t) \in C(S \otimes (t \geq 0)) \end{aligned}$$

y la condición de congruencia

$$\left.\left(\alpha u_0(x) + \beta \frac{\partial u_0(x)}{\partial \mathbf{n}}\right)\right|_{x \in S} = v(x,0), \quad x \in S$$

Para una ecuación parabólica, en particular la ecuación (8.33), el problema mixto se plantea de manera similar: hay que encontrar la función $u(x,t)$, $x = (x_1, x_2, \ldots, x_n) \in D$ y $t \geqslant 0$, de la clase $C^2(C_\infty) \cap C(\bar{C}_\infty)$, que satisface la ecuación (8.33) en el cilindro $C_\infty$, cumple la condición inicial, ecuación (8.34), y una condición en la frontera $S$ del dominio $D$ en cada instante $t$, ecuación (8.41).

## 8.4. Solución de ecuaciones diferenciales parciales

Dependiendo del tipo de EDP, existen diferentes métodos analíticos de su solución, mencionando entre ellos la solución por integración, transformadas integrales de Fourier y Laplace, separación de variables y funciones de Green. En esta sección se consideran tres métodos: el método de solución por integración, que es tanto simple como reducido; solución por la transformada de Laplace y el método de separación de variables.

### 8.4.1. Solución por integración

Siendo un método no muy general, lo explicaremos con ejemplos.

**Ejemplo 8.2.** *Para encontrar una solución $u = u(x, y)$ de la EDP de primer orden*

$$\frac{\partial u(x,y)}{\partial x} = 0, \quad -\infty < x, y < \infty$$

*ésta se integra con respecto a la variable $x$, considerando $y$ como un parámetro constante. De esta manera se obtiene la solución general $u = f(y)$, en donde $f(y)$ es una función cualquiera de la variable $y$.*

**Ejemplo 8.3.** *Dada la ecuación de segundo orden*

$$\frac{\partial^2 u(x,y)}{\partial x^2} = 0, \quad -\infty < x, y < \infty$$

*integrando una vez con respecto a $x$ lleva a $\partial u/\partial x = A(y)$; integrando por segunda vez, se tiene la solución general*

$$u(x,y) = A(y)\,x + B(y)$$

*donde $A(y)$ y $B(y)$ son funciones arbitrarias de $y$.*

**Ejemplo 8.4.** *La EDP*

$$\frac{\partial^2 u(x,y)}{\partial x \partial y} + \frac{\partial u(x,y)}{\partial y} = 1$$

*es lineal de segundo orden no homogénea. Efectuando el cambio*

$$\alpha(x,y) = \frac{\partial u(x,y)}{\partial y}$$

*la ecuación se escribe en la forma*

$$\frac{\partial \alpha(x,y)}{\partial x} + \alpha(x,y) = 1$$

*Ecuación que se resuelve por separación de variables, quedando*

$$\begin{aligned}\frac{\partial \alpha}{1-\alpha} &= \partial x \\ -\ln(1-\alpha) &= x + B(y) \\ \alpha(x,y) &= 1 - e^{-B(y)}e^{-x}\end{aligned}$$

*donde* $B(y)$ *es una función arbitraria de* $y$. *Integrando la ecuación* $\alpha(x,y) = \partial u/\partial y$, *se obtiene la solución*

$$u(x,y) = A(x) + y - e^{-x}\int e^{-B(y)}dy = A(x) + y - e^{-x}F(y)$$

*donde* $A(x)$ *y* $F(y) = \int e^{-B(y)}dy$ *son funciones arbitrarias.*

**Ejemplo 8.5.** *La EDP*

$$\frac{\partial^2 u(x,y)}{\partial x^2} - y^2 u(x,y) = 0$$

*es lineal homogénea de segundo orden. Considerando* $y$ *como parámetro constante, la ecuación característica de la ecuación diferencial ordinaria correspondiente es de la forma*

$$\lambda^2 - y^2 = 0, \quad \lambda_{1,2} = \pm y$$

*Por lo tanto, la solución es*

$$u(x,y) = A(y)e^{yx} + B(y)e^{-yx}$$

*en donde* $A(y)$ *y* $B(y)$ *son funciones arbitrarias de la variable* $y$. *Ésta no es la solución general, ya que, por ejemplo, la solución* $u = y^2e^{\pm x}$ *no se puede expresar de esta forma.*

De los ejemplos anteriores se ve que la solución por integración funciona principalmente para las EDP en forma de monomios con respecto a las derivadas parciales, esto es cuando la integración puede efectuarse en forma explícita.

### 8.4.2. Solución mediante la transformada de Laplace

En el capítulo 7 mostramos la aplicación de la transformada de Laplace para resolver ecuaciones diferenciales ordinarias. Esta transformada también puede aplicarse para resolver ecuaciones diferenciales parciales. En esta sección se explica la idea básica considerando dos ejemplos típicos. Recordamos que, en el caso de una EDO, la transformada de Laplace da lugar a una ecuación algebraica. Se verá que, en el caso de una EDP en dos variables independientes, el método proporciona una ecuación diferencial ordinaria en la transformada. Esto se debe a que se transforma la EDP dada con respecto a una de las variables independientes, por lo común $t$, de modo que las derivadas con respecto a la otra variable se quedan en la ecuación transformada. Si los coeficientes de la EDP dada no dependen de $t$, la transformación simplificará el problema.

### Una EDO de primer orden

Para resolver el problema

$$\begin{aligned} \frac{\partial w(x,t)}{\partial x} + x\frac{\partial w(x,t)}{\partial t} &= 0 \quad (t \geq 0) \\ w(x,0) &= 0, \quad w(0,t) = t \end{aligned} \tag{8.42}$$

se toma la transformada de Laplace de ésta con respecto a la variable $t$. Ya que

$$L\left\{\frac{\partial w(x,t)}{\partial x}\right\} = \int_0^\infty \frac{\partial w(x,t)}{\partial x} e^{-st}dt = \frac{\partial}{\partial x}\int_0^\infty w(x,t)\, e^{-st}dt = \frac{\partial}{\partial x}L\{w(x,t)\}$$

y

$$L\left\{\frac{\partial w(x,t)}{\partial t}\right\} = sL\{w(x,t)\} - w(x,0)$$

la ecuación transformada es

$$\frac{\partial W(x,s)}{\partial x} + xsW(x,s) = 0 \tag{8.43}$$

en donde $W(x,s) \equiv L\{w(x,t)\}$. Ésta puede considerarse como una ecuación diferencial ordinaria con respecto a $x$, ya que no tiene derivadas con respecto a $s$. La solución general de (8.43) es

$$W(x,s) = C(s)\exp\left(-sx^2/2\right)$$

Puesto que $L\{t\} = 1/s^2$, la condición $w(0,t) = t$ da $W(0,s) = 1/s^2$ y, por tanto,

$$W(0,s) = 1/s^2 = C(s)$$

De donde

$$W(x,s) = \frac{1}{s^2}\exp\left(-sx^2/2\right)$$

Ahora bien, $L^{-1}\left\{1/s^2\right\} = t$, de modo que la propiedad de desplazamiento en $t$, ecuación (7.22), da

$$\begin{aligned} w(x,t) &= L^{-1}\{W(x,s)\} \qquad (8.44) \\ &= \left(t - \frac{x^2}{2}\right)u_{x^2/2}(t) \\ &= \begin{cases} 0 & \text{si} \quad t < x^2/2 \\ \left(t - \frac{x^2}{2}\right) & \text{si} \quad t \geq x^2/2 \end{cases} \qquad (8.45) \end{aligned}$$

Como se procedió formalmente, tiene que verificarse que (8.44) satisface a (8.42).

### Propagación de onda en una cuerda semiinfinita

La propagación de una perturbación transversal en una cuerda elástica tensada se describe mediante la ecuación unidimensional de onda

$$\frac{\partial^2 w(x,t)}{\partial t^2} = a^2 \frac{\partial^2 w(x,t)}{\partial x^2} \tag{8.46}$$

en donde $w(x,t)$ es el desplazamiento transversal de la cuerda en el punto $x$ en el instante $t$. Suponemos que la cuerda es semiinfinita $(x \geq 0)$ y está en reposo sobre el eje $x$ para $t < 0$. Para el tiempo $t > 0$, el extremo izquierdo se mueve según la ley

$$w(0,t) = f(t) \tag{8.47}$$

en donde $f(t)$ es una función dada de $t$. Además, por la intuición física

$$\lim_{x\to\infty} w(x,t) = 0 \tag{8.48}$$

para cualquier instante dado $t \geq 0$. Según el planteamiento del problema, (8.47) y (8.48) son condiciones en la frontera. Las condiciones iniciales son

$$w(x,0) = 0, \quad \frac{\partial w(x,0)}{\partial t} = 0 \tag{8.49}$$

Para resolver dicho problema se toma la transformada de Laplace de (8.46) con respecto a $t$. Ahora bien,

$$L\left\{\frac{\partial^2 w(x,t)}{\partial t^2}\right\} = s^2 L\{w(x,t)\} - sw(x,0) - \frac{\partial w(x,0)}{\partial t} = s^2 L\{w(x,t)\}$$

$$L\left\{\frac{\partial^2 w(x,t)}{\partial x^2}\right\} = \frac{\partial^2}{\partial x^2} L\{w(x,t)\}$$

Así entonces, se obtiene la ecuación transformada

$$\frac{\partial^2 W(x,s)}{\partial x^2} = \frac{s^2}{a^2} W(x,s)$$

en donde $W(x,s) = L\{w(x,t)\}$. La ecuación transformada puede considerarse como una EDO para $W(x,s)$, considerando $x$ como la variable independiente y $s$ como parámetro. La solución general es

$$W(x,s) = A(s)\exp(sx/a) + B(s)\exp(-sx/a)$$

Para determinar las "constantes" $A(s)$ y $B(s)$ hay que usar condiciones en la frontera

$$W(0,s) = L\{w(0,t)\} = L\{f(t)\} \equiv F(s)$$

y

$$\lim_{x\to\infty} W(x,s) = 0 = \lim_{x\to\infty} \int_0^\infty w(x,t) e^{-st} dt = \int_0^\infty e^{-st} \lim_{x\to\infty} w(x,t) = 0$$

La última condición implica que $A(s) = 0$, porque $a > 0$ y $s > 0$. De donde se tiene

$$B(s) = W(0, s) = F(s)$$

de modo que queda

$$W(x, s) = F(s) \exp(-sx/a)$$

La transformada inversa queda

$$\begin{aligned} w(x,t) &= L^{-1}\{W(x,s)\} \quad (8.50) \\ &= f\left(t - \frac{x}{a}\right) u_{x/a}(t) \\ &= \begin{cases} 0 & \text{si} \quad t < x/a \\ f\left(t - \frac{x}{a}\right) & \text{si} \quad t \geq x/a \end{cases} \quad (8.51) \end{aligned}$$

Nótese que un punto $x$ permanece en reposo hasta $t = x/a$, el tiempo necesario para que la perturbación, que viaja con la velocidad $a$, llegaría hasta esa $x$. Dado que se procedió de manera formal, debe verificarse que (8.50) realmente satisface tanto la ecuación de onda como las condiciones en la frontera e iniciales.

## 8.4.3. El método de Fourier (separación de variables)

La idea básica del **método de Fourier** (**separación de variables**) consiste en buscar la solución en forma del producto de funciones, algunas (o en el mejor de los casos, cada una) de éstas siendo función de una sola variable independiente. El método es muy efectivo para la solución de problemas con valores frontera. En los ejemplos que siguen se explican los principios del método de separación de variables para las ecuaciones de onda y conductividad térmica, concluyendo con una generalización para los problemas de valores frontera para las ecuaciones de diferentes tipos.

### Separación de variables para la ecuación de onda

Se conoce que las vibraciones pequeñas de una cuerda elástica se rigen por la ecuación unidimensional de onda

$$\frac{\partial^2 u(x,t)}{\partial t^2} = a^2 \frac{\partial^2 u(x,t)}{\partial x^2}, \quad 0 < x < L, \quad t > 0 \qquad (8.52)$$

donde $u(x,t)$ es la deflexión de la cuerda. Suponemos que los dos extremos de la cuerda son fijos, es decir, las condiciones en la frontera son

$$u(0,t) = 0, \quad u(L,t) = 0 \qquad (8.53)$$

Además, suponemos que las dos condiciones iniciales son

$$u(x,0) = a(x), \quad u_t(x,0) = b(x) \qquad (8.54)$$

donde $a(x)$ y $b(x)$ son funciones dadas de $x$.

Se considera la solución de dicha ecuación por el método de separación de variables, es decir, se busca una solución del problema con valores frontera de la forma

$$u(x,t) = X(x)\,T(t) \tag{8.55}$$

Al sustituir esta expresión en la ecuación (8.52) y dividiendo entre $a^2XT$, se tiene

$$\frac{X''}{X} = \frac{\ddot{T}}{a^2T}$$

donde la prima denota la derivada con respecto a $x$ y el punto la derivada con respecto a $t$. La expresión de la izquierda comprende funciones que sólo dependen de $x$, en tanto que la de la derecha contiene funciones que dependen sólo de $t$. De donde, la igualdad anterior puede ser cierta solamente si ambas expresiones tienen el mismo valor constante, digamos $-\lambda$,

$$\frac{X''}{X} = \frac{\ddot{T}}{a^2T} = -\lambda$$

De aquí

$$X'' + \lambda X = 0 \tag{8.56}$$

$$\ddot{T} + a^2\lambda T = 0 \tag{8.57}$$

La solución (8.55) debe satisfacer las condiciones en la frontera, ecuación (8.54). Por lo tanto, la función $X(x)$ debe satisfacer las condiciones

$$X(0) = X(L) = 0 \tag{8.58}$$

Se distinguen dos casos. Cuando $\lambda \leq 0$, la única solución para la ecuación (8.56), que satisface las condiciones (8.58), es $X(x) = 0$. En el caso de $\lambda > 0$, se puede hacer $\lambda = k^2$, con lo que la ecuación (8.56) da como resultado

$$X(x) = A\cos kx + B\,\mathrm{sen}\,kx$$

La condición $X(0) = 0$ exige $A = 0$, y así queda $X(x) = B\,\mathrm{sen}\,kx$. Debe tomarse $B \neq 0$, de lo contrario $X(x) \equiv 0$. Por lo tanto, de la condición $X(L) = 0$ sigue que $\mathrm{sen}\,kL = 0$. En consecuencia, $kL = n\pi$ y $k$ puede tomar cualquiera de los valores

$$k_n = \frac{n\pi}{L}, \quad n = 1, 2, \ldots \tag{8.59}$$

Entonces, se obtiene una infinidad de soluciones

$$X_n(x) = \mathrm{sen}\,\frac{n\pi}{L}x, \quad n = 1, 2, \ldots \tag{8.60}$$

que reciben el nombre de **modos normales**. Para los valores permitidos $k_n$, de la ecuación

$$\ddot{T} + a^2k_n^2T = 0$$

tenemos

$$T_n(t) = C_n \cos \omega_n t + D_n \operatorname{sen} \omega_n t \tag{8.61}$$

donde $C_n$ y $D_n$ son constantes arbitrarias y $\omega_n = ak_n$ es la frecuencia (angular) normal del modo $X_n(x)$ correspondiente. Las funciones

$$u_n(x,t) = X_n(x) T_n(t) = \operatorname{sen} k_n x \left(C_n \cos \omega_n t + D_n \operatorname{sen} \omega_n t\right) \tag{8.62}$$

son soluciones de la ecuación (8.52) que satisfacen las condiciones en la frontera (8.53), si no, en general, no satisfacen las condiciones iniciales. Estas funciones se conocen como **eigenfunciones** (funciones características) del problema correspondiente con valores frontera, y los valores $\omega_n = ak_n = an\pi/L$ se llaman **eigenvalores** (valores característicos) de la cuerda vibrante. El conjunto $\{\omega_1, \omega_2, \ldots\}$ se llama **espectro**.

Tanto la ecuación de onda (8.52) como las condiciones en la frontera (8.53) son lineales y homogéneas. Por consiguiente, cualquier combinación lineal de soluciones (8.62) es una solución de la ecuación (8.52) que satisface la ecuación (8.53). Por lo tanto, para satisfacer las condiciones iniciales, se buscará ahora la solución del problema con valores frontera, ecuaciones (8.52), (8.53) y (8.54), en la forma de serie infinita

$$u(x,t) = \sum_{n=1}^{\infty} \operatorname{sen} k_n x \left(C_n \cos \omega_n t + D_n \operatorname{sen} \omega_n t\right) \tag{8.63}$$

Asumiendo que la serie es convergente y diferenciable término a término dos veces con respecto a $x$ y a $t$, se encuentra de las condiciones iniciales que

$$u(x,0) = a(x) = \sum_{n=1}^{\infty} C_n \operatorname{sen} k_n x \tag{8.64}$$

$$u_t(x,0) = b(x) = \sum_{n=1}^{\infty} \omega_n D_n \operatorname{sen} k_n x \tag{8.65}$$

donde $\omega_n = ak_n = an\pi/L$. Se reconocen esas expresiones como series senoidales de Fourier de las funciones $a(x)$ y $b(x)$ en el intervalo $(0, L)$. Por lo tanto, los coeficientes están dados por las siguientes ecuaciones

$$C_n = \frac{2}{L} \int_0^L a(x) \operatorname{sen}(k_n x)\, dx \tag{8.66}$$

$$\omega_n D_n = \frac{2}{L} \int_0^L b(x) \operatorname{sen}(k_n x)\, dx \tag{8.67}$$

Finalmente, sustituyendo $C_n$ y $D_n$ en la ecuación (8.63), se obtiene la solución de la ecuación (8.52) que satisface las condiciones (8.53) y (8.54).

Notamos que los mismos modos normales pueden aplicarse a la ecuación de onda de la forma más general. Se considera la ecuación ($\alpha$ y $\beta$ son constantes)

$$u_{tt} + \alpha u_t + \beta u = a^2 u_{xx} \tag{8.68}$$

sujeta a las condiciones en la frontera, ecuación (8.53). Las ecuaciones para $X(x)$ y $T(t)$, en el método de separación de variables, son

$$\begin{aligned} X'' + k^2 X &= 0 \\ \ddot{T} + \alpha \dot{T} + (\beta + a^2 k^2) T &= 0 \end{aligned} \tag{8.69}$$

Por lo tanto, los modos normales son otra vez $X_n(x) = \operatorname{sen}(n\pi x/L)$. Las soluciones normales, $u_n(x,t) = X_n(x) T_n(t)$, sin embargo, son modificadas porque $T(t)$ ahora tiene que ser una solución de la ecuación (8.69).

**Flujo de calor en una barra delgada**

El flujo de calor en un medio isotrópico y homogéneo se describe por la ecuación de conductividad térmica, ecuación (8.22). Como una aplicación importante, se considera el caso de una varilla larga y delgada de sección constante, la cual está orientada a lo largo del eje $x$. Se supone que la varilla se encuentra perfectamente aislada en toda su superficie lateral, de modo que el calor sólo fluye en la dirección $x$. Entonces, la temperatura $u$ dependerá nada más de $x$ y del tiempo $t$, y la ecuación de conductividad térmica se convierte en la ecuación unidimensional

$$\frac{\partial u(x,t)}{\partial t} = a^2 \frac{\partial^2 u(x,t)}{\partial x^2} \tag{8.70}$$

donde $a^2 = p/c\tilde{\rho}$, $p$ es la conductividad térmica, $c$ es el calor específico y $\tilde{\rho}$ es la densidad del material del medio. Se supone que los extremos $x = 0$ y $x = L$ de la varilla se mantienen a la temperatura cero, es decir, las condiciones en la frontera son

$$u(0,t) = 0, \quad u(L,t) = 0 \tag{8.71}$$

Nótese que éstas tienen la misma forma que las (8.53) para la ecuación unidimensional de onda (8.52). Sea $a(x)$ la temperatura inicial de la varilla; por lo tanto, la condición inicial es

$$u(x,0) = a(x) \tag{8.72}$$

en donde $a(x)$ es una función dada. El procedimiento para resolver la ecuación (8.70) con dichos valores frontera será semejante al aplicado en el caso de la ecuación de onda. Sin embargo, el comportamiento de las soluciones será completamente diferente al de las correspondientes de la ecuación de onda, debido a que la ecuación (8.70) comprende a $u_t$, en tanto que la ecuación de onda contiene a $u_{tt}$.

Según el método de separación de variables, se busca una solución de la forma $u(x,t) = X(x) T(t)$. Sustituyendo esta expresión en la ecuación (8.70) y dividiendo el resultado entre $a^2 XT$, se tiene

$$\frac{\dot{T}}{a^2 T} = \frac{X''}{X} \tag{8.73}$$

La expresión de la izquierda depende sólo de $t$, mientras que la de la derecha sólo depende de $x$. De donde se concluye que las dos expresiones deben ser

iguales a una constante, digamos $-\lambda$. Si $\lambda \leq 0$, la única solución $u = XT$ que satisface las condiciones en la frontera, ecuación (8.71), es $u \equiv 0$. Para $\lambda > 0$, tomamos $\lambda = k^2$ y de la ecuación (8.73) se obtienen dos ecuaciones

$$X'' + k^2 X = 0 \tag{8.74}$$
$$\dot{T} + a^2 k^2 T = 0 \tag{8.75}$$

La solución general de la ecuación (8.74) es

$$X(x) = A \cos kx + B \operatorname{sen} kx$$

De las condiciones en la frontera se deduce que la función $X(x)$ debe satisfacer las condiciones $X(0) = 0$ y $X(L) = 0$. La condición $X(0) = 0$ exige $A = 0$, y así queda $X(x) = B \operatorname{sen} kx$. Debe tomarse $B \neq 0$, de lo contrario $X(x) \equiv 0$. Por tanto, de la condición $X(L) = 0$ se sigue que $\operatorname{sen} kL = 0$. En consecuencia, $kL = n\pi$ y $k$ puede tomar cualquiera de los valores

$$k_n = \frac{n\pi}{L}, \quad n = 1, 2, \ldots \tag{8.76}$$

Entonces, se obtiene una infinidad de soluciones

$$X_n(x) = \operatorname{sen} \frac{n\pi}{L} x, \quad n = 1, 2, \ldots \tag{8.77}$$

Para los valores permitidos $k_n$, de la ecuación

$$\dot{T} + a^2 k_n^2 T = 0$$

tenemos

$$T_n(t) = C_n \exp\left(-\omega_n^2 t\right) \tag{8.78}$$

donde $\omega_n = a k_n$ y $C_n$ es una constante arbitraria. Las funciones

$$u_n(x,t) = X_n(x) T_n(t) = C_n \operatorname{sen} k_n x \exp\left(-\omega_n^2 t\right) \tag{8.79}$$

son soluciones de la ecuación (8.70) que satisfacen las condiciones en la frontera (8.71), si no, no satisfacen las condiciones iniciales.

Tanto la ecuación de conductividad térmica (8.70) como las condiciones en la frontera (8.71) son lineales y homogéneas. Para obtener una solución que también satisfaga la condición inicial, se considera la serie infinita

$$u(x,t) = \sum_{n=1}^{\infty} u_n(x,t) = \sum_{n=1}^{\infty} C_n \operatorname{sen} k_n x \exp\left(-\omega_n^2 t\right) \tag{8.80}$$

Asumiendo que la serie es convergente y diferenciable término a término dos veces con respecto a $x$ y una vez con respecto a $t$, de la ecuación (8.72) se encuentra que

$$u(x,0) = a(x) = \sum_{n=1}^{\infty} C_n \operatorname{sen} k_n x \tag{8.81}$$

Se considera esa expresión como una serie senoidal de Fourier de $a(x)$ en el intervalo $(0, L)$. Por lo tanto, los coeficientes están dados por la siguiente ecuación

$$C_n = \frac{2}{L}\int_0^L a(x)\,\mathrm{sen}\,(k_n x)\,dx \qquad (8.82)$$

Finalmente, sustituyendo $C_n$ en la ecuación (8.80), se obtiene la solución de la ecuación (8.70) que satisface las condiciones (8.71) y (8.72). Nótese que la solución (8.80) es completamente diferente de la solución (8.63) para la ecuación de onda. Debido al factor exponencial, todos los términos de (8.80) tienden a cero cuando $t$ tiende al infinito. La rapidez del decremento varía con $n$.

### 8.4.4. Separación de variables en problemas con valores frontera

Una generalización del método de Fourier (separación de variables) para las ecuaciones del tipo hiperbólico y parabólico se considera en esta sección.

**Propiedades del operador $\hat{L}$**

Sea $\mathcal{L}_2(D,\rho)$ un espacio lineal de las funciones $u(x)$, $x = (x_1, x_2, \ldots, x_n) \in D$, tales que la función $\rho(x)\,|u(x)|^2$ es integrable en $D$, donde $\rho(x) \in C(\bar{D})$, $\rho(x) > 0$, es una función peso y $\bar{D} = D \cup S$, siendo $S$ la frontera de $D$. Al introducir el producto escalar y la norma en este espacio según las fórmulas

$$(u,v)_\rho = \int_D \rho(x)\,u(x)\,v^*(x)\,dx, \quad \|u\|_\rho = \sqrt{(u,u)_\rho} \qquad (8.83)$$

lo convierte en un espacio normalizado.

En el espacio $\mathcal{L}_2(D,\rho)$ definimos un operador diferencial lineal $\hat{L}$ por medio de la expresión

$$\begin{aligned} \hat{L}u &= -\operatorname{div}(p\,\mathrm{grad}\,u) + qu \qquad (8.84)\\ p(x) &\in C^1(\bar{D}), \quad q(x) \in C(\bar{D}) \\ p(x) &> 0, \quad q(x) \geq 0 \text{ para } x \in \bar{D} \end{aligned}$$

y la condición en la frontera $S$

$$\begin{aligned} \left(\alpha u + \beta \frac{\partial u}{\partial \mathbf{n}}\right)\Big|_S &= 0 \qquad (8.85)\\ \alpha(x), \beta(x) &\in C(S) \\ \alpha + \beta &> 0, \quad \alpha(x) \geq 0, \quad \beta(x) \geq 0 \text{ para } x \in S \end{aligned}$$

Es decir, el dominio de definición $G_L$ del operador $\hat{L}$ está constituido por todas las funciones $u(x) \in \mathcal{L}_2(D,\rho)$ que satisfacen la ecuación (8.85) y $\hat{L}u \in \mathcal{L}_2(D,\rho)$.

El operador $\hat{L}$ tiene aplicaciones importantes en relación con el método de separación de variables. Por lo tanto, a continuación se consideran (sin demostraciones) propiedades fundamentales relacionadas con el problema de eigenvalores de este operador.

De por sí, el problema de eigenvalores se plantea de la siguiente manera. Hay que encontrar los valores del parámetro $\lambda$ (eigenvalores del operador $\hat{L}$) tales que la ecuación

$$\hat{L}u = \lambda \rho u \tag{8.86}$$

tiene soluciones no triviales (no iguales a cero) en el dominio $G_L$ con una función peso $\rho(x) > 0$. Estas funciones son eigenfunciones de $\hat{L}$.

Es fácil ver que el operador $\hat{L}$ es lineal, cumpliéndose

$$\hat{L}(u_1 + u_2) = \hat{L}u_1 + \hat{L}u_2 \tag{8.87}$$

Aplicando las fórmulas de Green y Gauss-Ostrogradsky puede demostrarse que el operador $\hat{L}$ es hermitiano autoadjunto (véase, por ejemplo, [18])

$$\left(\hat{L}u, v\right) = \left(u, \hat{L}v\right) \text{ para } u, v \in G_L \tag{8.88}$$

y, además, definido positivo

$$\left(\hat{L}u, u\right) \geq 0 \text{ para } u \in G_L \tag{8.89}$$

donde el producto escalar $(u, v) = \int u(x) v^*(x) dx$. En consecuencia, los eigenvalores de $\hat{L}$ son no negativos. Sean $\lambda_k$ $(k = 1, 2, ...)$ los eigenvalores de $\hat{L}$ numerados en orden creciente de sus magnitudes: $0 \leq \lambda_1 \leq \lambda_2 \leq \cdots . \leq \lambda_k \leq \cdots$, repitiéndose cada $\lambda_k$ un número de veces igual a la multiplicidad de este eigenvalor. Sean $X_1, X_2, \ldots, X_k, \ldots$ eigenfunciones correspondientes, de tal manera que a cada $\lambda_k$ le corresponde una y sólo una eigenfunción $X_k$,

$$\hat{L}X_k = \lambda_k \rho X_k, \quad X_k \in G_L, \quad k = 1, 2, \ldots \tag{8.90}$$

Las eigenfunciones que corresponden a diferentes eigenvalores son ortogonales. Las eigenfunciones de $\hat{L}$ pueden escogerse reales y ortonormales

$$\begin{aligned} \left(\hat{L}X_k, X_l\right) &= \lambda_k (X_k, X_l)_\rho = \lambda_k \delta_{kl} \\ (X_k, X_l)_\rho &= \int_D \rho(x) X_k(x) X_l(x) dx \end{aligned} \tag{8.91}$$

Además, el conjunto de eigenfunciones de $\hat{L}$ es completo en $\mathcal{L}_2(D, \rho)$ y cada función $u(x) \in G_L$ puede ser presentada en forma de serie

$$u(x) = \sum_{k=1}^{\infty} (u, X_k)_\rho X_k(x) \tag{8.92}$$

La última fórmula es fundamental para la aplicación del método de separación de variables, como se ve a continuación.

Los resultados anteriores, dados aquí sin demostraciones, se pueden resumir mediante el siguiente teorema.

**Teorema 8.1.** *El conjunto de eigenvalores del operador $\hat{L}$ es contable y este conjunto no tiene un punto límite finito; cada eigenvalor tiene multiplicidad finita. Cada función del espacio lineal $G_L$ puede ser representada en forma de serie de Fourier por eigenfunciones del operador $\hat{L}$.*

### Ecuación hiperbólica homogénea

En el caso de un problema mixto para la ecuación de onda

$$\rho(x)\frac{\partial^2 u}{\partial t^2} = -\hat{L}u, \qquad u = u(x,t) \tag{8.93}$$

que es del tipo hiperbólico, las condiciones iniciales son

$$u(x,0) = a(x), \quad u_t(x,0) = b(x), \quad x \in D \tag{8.94}$$

y las condiciones en la frontera

$$\left.\left(\alpha u + \beta\frac{\partial u}{\partial \mathbf{n}}\right)\right|_S = 0, \quad t > 0 \tag{8.95}$$

El método de Fourier (separación de variables) consiste principalmente en lo siguiente. Se buscan las soluciones de la ecuación (8.93) en la forma de producto $X(x)T(t)$ que satisfacen la condición en la frontera, ecuación (8.95). De estas soluciones se compone una combinación lineal que satisface las condiciones iniciales, siendo ésta la solución del problema con valores en la frontera, ecuaciones (8.93), (8.94) y (8.95).

Entonces, por separación de variables, se busca una solución $u(x,t) = X(x)T(t)$, que cumple la condición en la frontera (8.95). Sustituyéndola en la ecuación (8.93) y dividiendo entre $\rho XT$, se obtiene

$$\frac{\ddot{T}(t)}{T(t)} = \frac{\hat{L}X(x)}{\rho(x)X(x)} \tag{8.96}$$

La parte izquierda de la igualdad podría depender solamente de $t$, en tanto que la derecha podría depender de $x$. Por tanto, para que se cumpla la igualdad, ambas partes deben ser iguales a una constante. Denominando esta constante como $-\lambda$, de la ecuación (8.96) se llega a dos ecuaciones:

$$\hat{L}X = \lambda\rho(x)X \tag{8.97}$$

$$\ddot{T} + \lambda T = 0 \tag{8.98}$$

La ecuación (8.93) se partió en las dos, ecuaciones (8.97) y (8.98), cada una de un número menor de variables independientes. En esta ocasión se dice que las variables se separaron.

La solución $X(x)$ de la ecuación (8.97) debe satisfacer la condición en la frontera

$$\left.\left(\alpha X + \beta\frac{\partial X}{\partial \mathbf{n}}\right)\right|_S = 0 \tag{8.99}$$

Por lo tanto, estas soluciones son eigenfunciones $X_k(x)$ con eigenvalores correspondientes $\lambda_k$ del operador $\hat{L}$. Para $\lambda = \lambda_k$ una solución general de la ecuación (8.98) es

$$T_k(t) = A_k \cos\sqrt{\lambda_k}t + B_k \operatorname{sen}\sqrt{\lambda_k}t, \quad k = 1, 2, \ldots \tag{8.100}$$

donde $A_k$ y $B_k$ son constantes arbitrarias. Cada función

$$u_k(x,t) = X_k(x) T_k(t) = X_k(x)\left(A_k \cos\sqrt{\lambda_k}t + B_k \operatorname{sen}\sqrt{\lambda_k}t\right) \tag{8.101}$$

es una solución particular de la ecuación (8.93) que satisface la condición en la frontera, ecuación (8.95). Una combinación lineal de estas soluciones es también una solución que satisface la ecuación (8.93) con la condición en la frontera (8.95). Por lo tanto, como un candidato para la solución del problema se considera una serie formal

$$u(x,t) = \sum_{k=1}^{\infty} X_k(x) T_k(t) = \sum_{k=1}^{\infty} X_k(x)\left(A_k \cos\sqrt{\lambda_k}t + B_k \operatorname{sen}\sqrt{\lambda_k}t\right) \tag{8.102}$$

Los coeficientes $A_k$ y $B_k$ se escogen de tal manera que la serie satisface las condiciones iniciales dadas por la ecuación (8.94)

$$u(x,0) = a(x) = \sum_{k=1}^{\infty} A_k X_k(x), \quad u_t(x,0) = b(x) = \sum_{k=1}^{\infty} \sqrt{\lambda_k} B_k X_k(x)$$

es decir,

$$A_k = (a, X_k)_\rho = \int_D \rho(x)\, a(x)\, X_k(x)\, dx, \quad B_k = \frac{1}{\sqrt{\lambda_k}} (b, X_k)_\rho \tag{8.103}$$

dado que el conjunto de eigenfunciones $\{X_k(x)\}$ es completo en $\mathcal{L}_2(D,\rho)$. Entonces, la serie (8.102) con coeficientes $A_k$ y $B_k$ dados por la ecuación (8.103) es la solución del problema mixto, ecuaciones (8.93), (8.94) y (8.95). El miembro $k$-ésimo de la serie (8.102)

$$X_k(x) T_k(t) = C_k X_k(x) \operatorname{sen}\left(\sqrt{\lambda_k}t + \alpha_k\right)$$

donde

$$C_k = \sqrt{A_k^2 + B_k^2}, \quad \operatorname{sen}\alpha_k = A_k/C_k, \quad \cos\alpha_k = B_k/C_k$$

es una oscilación armónica de frecuencia $\sqrt{\lambda_k}$ y de amplitud $C_k X_k(x)$.

### Ecuación hiperbólica no homogénea

En esta sección presentamos una versión más general del método de Fourier (separación de variables) que sirve para la solución del problema mixto para la **ecuación hiperbólica no homogénea**

$$\rho(x) \frac{\partial^2 u}{\partial t^2} = -\hat{L}u + F(x,t) \tag{8.104}$$

Para $t > 0$ presentamos la solución del problema, planteado por las ecuaciones (8.104), (8.94) y (8.95), en forma de serie de Fourier por eigenfunciones $\{X_k\}$ del operador $\hat{L}$

$$u(x,t) = \sum_{k=1}^{\infty} X_k(x)\, T_k(t), \quad T_k(t) = (u, X_k)_\rho \tag{8.105}$$

Con el fin de encontrar la ecuación diferencial para las funciones $T_k(t)$, la solución (8.105) se sustituye en la ecuación (8.104)

$$\begin{aligned} \sum_{m=1}^{\infty} \rho(x)\, X_m(x)\, \ddot{T}_m(t) &= -\sum_{m=1}^{\infty} T_m(t)\, \hat{L} X_m(x) + F(x,t) \\ &= -\sum_{m=1}^{\infty} T_m(t)\, \lambda_m \rho X_m(x) + F(x,t) \end{aligned}$$

Después se toma el producto escalar de esta ecuación por la eigenfunción $X_k$,

$$\sum_{m=1}^{\infty} (X_m, X_k)_\rho\, \ddot{T}_m(t) = -\sum_{m=1}^{\infty} T_m(t)\, \lambda_m (X_m, X_k)_\rho + (F, X_k)$$

y, usando la ortonormalidad de eigenfunciones, se obtienen las ecuaciones

$$\ddot{T}_k(t) + \lambda_k T_k(t) = f_k(t), \quad k = 1, 2, \ldots \tag{8.106}$$

$$f_k(t) = (F, X_k) = \int_D F(x,t)\, X_k(x)\, dx \tag{8.107}$$

Debido a las condiciones iniciales, ecuaciones (8.94) y (8.103), las funciones incógnitas $T_k(t)$ deben satisfacer las siguientes condiciones iniciales:

$$T_k(0) = \int_D \rho(x)\, a(x)\, X_k(x)\, dx = (a, X_k)_\rho = A_k \tag{8.108}$$

$$\begin{aligned} \dot{T}_k(0) &= (u_t(x,0), X_k)_\rho \\ &= \int_D \rho(x)\, b(x)\, X_k(x)\, dx = (b, X_k)_\rho = \sqrt{\lambda_k} B_k \end{aligned} \tag{8.109}$$

La solución del problema de Cauchy para la ecuación (8.106) con condiciones iniciales, ecuaciones (8.108) y (8.109), es

$$\begin{aligned} T_k(t) &= A_k \cos\sqrt{\lambda_k} t + B_k \operatorname{sen}\sqrt{\lambda_k} t \\ &\quad + \frac{1}{\sqrt{\lambda_k}} \int_0^t f_k(\tau) \operatorname{sen}\sqrt{\lambda_k}(t-\tau)\, d\tau \end{aligned} \tag{8.110}$$

Sustituyendo esta expresión en la serie (8.105), se obtiene la solución formal del problema mixto, ecuaciones (8.94), (8.95) y (8.104),

$$u(x,t) = \sum_{k=1}^{\infty} X_k(x) \left[ \begin{array}{l} A_k \cos\sqrt{\lambda_k} t + B_k \operatorname{sen}\sqrt{\lambda_k} t \\ + \frac{1}{\sqrt{\lambda_k}} \int_0^t f_k(\tau) \operatorname{sen}\sqrt{\lambda_k}(t-\tau)\, d\tau \end{array} \right] \tag{8.111}$$

Nótese que los dos primeros términos de la serie representan la solución del problema mixto para la ecuación homogénea con $F(x,t) \equiv 0$; el tercer término es la solución del problema mixto para la ecuación no homogénea con condiciones iniciales $a(x) = b(x) = 0$.

Sea $a(x) = b(x) = 0$ y

$$F(x,t) = C\rho(x) X_m(x) \operatorname{sen} \omega t \tag{8.112}$$

Entonces,

$$A_k = B_k = 0, \quad f_k(t) = C(X_m, X_k)_\rho \operatorname{sen} \omega t = C\delta_{mk} \operatorname{sen} \omega t$$

y, en consecuencia, de la ecuación (8.110) se tiene

$$\begin{aligned} T_k(t) &= \frac{C\delta_{mk}}{\sqrt{\lambda_k}} \int_0^t \operatorname{sen} \omega\tau \operatorname{sen} \sqrt{\lambda_k}(t-\tau)\, d\tau \\ &= \frac{C\delta_{mk}}{\omega^2 - \lambda_m} \left( \frac{\omega}{\sqrt{\lambda_m}} \operatorname{sen} \sqrt{\lambda_m} t - \operatorname{sen} \omega t \right) \end{aligned}$$

Por tanto, la serie (8.111) se reduce a un solo término

$$u(x,t) = \frac{C}{\omega^2 - \lambda_m} \left( \frac{\omega}{\sqrt{\lambda_m}} \operatorname{sen} \sqrt{\lambda_m} t - \operatorname{sen} \omega t \right) X_m(x) \tag{8.113}$$

que es la solución del problema. Cuando $\omega \to \sqrt{\lambda_m}$, la solución (8.113) toma la forma

$$u(x,t) = \frac{C}{2\sqrt{\lambda_m}} \left( \frac{\operatorname{sen} \sqrt{\lambda_m} t}{\sqrt{\lambda_m}} - t \cos \sqrt{\lambda_m} t \right) X_m(x) \tag{8.114}$$

La última fórmula indica que cuando un sistema vibrante sin amortiguamiento está sometido a la acción de una fuerza externa (8.112) de frecuencia igual a una de las frecuencias propias, $\sqrt{\lambda_m}$, la amplitud de las oscilaciones de este sistema crece hasta el infinito cuando $t \to \infty$; este fenómeno se conoce como **resonancia**.

### Ecuación parabólica no homogénea

En esta sección se considera la solución de un problema mixto para la **ecuación parabólica no homogénea**

$$\rho(x) \frac{\partial u}{\partial t} = -\hat{L}u + F(x,t) \tag{8.115}$$

con la condición inicial

$$u(x,0) = a(x), \quad x \in D \tag{8.116}$$

y condiciones en la frontera

$$\left( \alpha u + \beta \frac{\partial u}{\partial \mathbf{n}} \right) \Bigg|_S = 0, \quad t > 0 \tag{8.117}$$

De acuerdo con el método de separación de variables, la solución $u(x,t)$ se busca en forma de la serie (8.105). Para las funciones $T_k(t)$ se obtiene el siguiente problema de Cauchy:

$$\dot{T}_k(t) + \lambda_k T_k(t) = f_k(t), \quad T_k(0) = (a, X_k)_\rho = A_k, \quad k = 1, 2, \dots \tag{8.118}$$

donde $f_k(t)$ y $A_k$ se definen mediante las ecuaciones (8.107) y (8.108), respectivamente. La solución del problema (8.118) es

$$T_k(t) = A_k e^{-\lambda_k t} + \int_0^t f_k(\tau) e^{-\lambda_k (t-\tau)} d\tau \tag{8.119}$$

Por lo tanto, la solución formal del problema mixto, ecuaciones (8.115), (8.116) y (8.117), se da mediante la serie

$$u(x,t) = \sum_{k=1}^{\infty} X_k(x) \left[ A_k e^{-\lambda_k t} + \int_0^t f_k(\tau) e^{-\lambda_k (t-\tau)} d\tau \right] \tag{8.120}$$

**Ecuación de Schrodinger**

El problema mixto para la ecuación de Schrodinger, que es homogénea del tipo parabólico,

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \Delta \psi + V(x) \psi \tag{8.121}$$

$$\psi(x,0) = \psi_0(x), \quad x \in D \tag{8.122}$$

$$\left. \left( \alpha \psi + \beta \frac{\partial \psi}{\partial \mathbf{n}} \right) \right|_S = 0, \quad t > 0 \tag{8.123}$$

se considera de manera similar al problema anterior, ecuaciones (8.115), (8.116) y (8.117). Para las funciones $T_k(t)$ se obtiene el siguiente problema de Cauchy:

$$i\hbar \dot{T}_k(t) + \lambda_k T_k(t) = 0, \quad T_k(0) = (\psi_0, X_k) = A_k, \quad k = 1, 2, \dots \tag{8.124}$$

donde $\{X_k(x)\}$ son eigenfunciones del operador $\hat{L}$ con $p = \hbar^2/2m$ y $q = V(x)$. La solución de la ecuación (8.124) es

$$T_k(t) = A_k e^{-i\lambda_k t/\hbar} \tag{8.125}$$

y, en consecuencia, la solución del problema mixto para la ecuación de Schrodinger se puede expresar en forma de la serie

$$\psi(x,t) = \sum_{k=1}^{\infty} A_k e^{-i\lambda_k t/\hbar} X_k(x) \tag{8.126}$$

**Nota bibliográfica:** para ampliar los conocimientos sobre el tema de este capítulo, puede consultar [1], [2], [4], [5], [6], [11], [17] y [18].

De acuerdo con el método de separación de variables, la solución se busca en forma de la serie (8.105). Para las funciones $T_k(t)$ se obtiene el siguiente problema de Cauchy:

$$T_k'(t) + \lambda_k T_k(t) = f_k(t), \quad T_k(0) = (\varphi, X_k)_\rho = A_k, \quad k = 1, 2, \ldots$$

donde $f_k(t)$ y $A_k$ se definen mediante las ecuaciones (8.107) y (8.108), respectivamente. La solución del problema (8.118) es

$$T_k(t) = A_k e^{-\lambda_k t} + \int_0^t f_k(\tau) e^{-\lambda_k (t-\tau)} d\tau$$

Por lo tanto, la solución formal del problema mixto, ecuaciones (8.115), (8.116) y (8.117), se da mediante la serie

$$u(x,t) = \sum_{k=1}^{\infty} X_k(x) \left[ A_k e^{-\lambda_k t} + \int_0^t f_k(\tau) e^{-\lambda_k (t-\tau)} d\tau \right]$$

## Ecuación de Schrödinger

El problema mixto para la ecuación de Schrödinger, que es homogénea del tipo parabólico,

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m}\Delta \psi + V(x)\psi$$

$$\psi(x,0) = \psi_0(x), \quad x \in D$$

$$\left( \alpha\psi + \beta \frac{\partial \psi}{\partial n} \right)\bigg|_S = 0, \quad t > 0$$

se considera de manera similar al problema anterior, ecuaciones (8.115), (8.116) y (8.117). Para las funciones $T_k(t)$ se obtiene el siguiente problema de Cauchy:

$$i\hbar T_k'(t) + \lambda_k T_k(t) = 0, \quad T_k(0) = (\psi_0, X_k) = A_k, \quad k = 1, 2, \ldots$$

donde $\{X_k(x)\}$ son eigenfunciones del operador $L$ con $p = \hbar^2/2m$ y $q = V$. La solución de la ecuación (8.124) es

$$T_k(t) = A_k e^{-i\lambda_k t/\hbar}$$

y, en consecuencia, la solución del problema mixto para la ecuación de Schrödinger se puede expresar en forma de la serie

$$\psi(x,t) = \sum_{k=1}^{\infty} A_k e^{-i\lambda_k t/\hbar} X_k(x)$$

**Nota bibliográfica:** para ampliar los conocimientos sobre el tema de este capítulo, puede consultar [1], [2], [3], [4], [5], [6], [11], [17] y [18].

# Capítulo 9

# Teoría de Sturm-Liouville y funciones especiales

En el capítulo anterior se vio que el método de separación de variables en problemas con valores frontera nos lleva al problema de eigenvalores para el operador $\hat{L}$, ecuaciones (8.85) y (8.86). Un caso particular, pero sumamente importante en las Matemáticas aplicadas a la Física e Ingeniería, es el problema de eigenvalores para el operador $\hat{L}$ unidimensional, es decir, de una sola variable independiente. Se verá que las funciones especiales de Bessel, Legendre, Hermite y otras están relacionadas estrechamente con el problema de Sturm-Liouville y surgen como eigenfunciones del operador unidimensional $\hat{L}$.

## 9.1. Problema de Sturm-Liouville

El problema de eigenvalores para el operador $\hat{L}$, ecuaciones (8.85) y (8.86), en el caso de una sola variable independiente $x$, se llama **problema de Sturm-Liouville**

$$\hat{L}u \equiv -(pu')' + qu = \lambda \rho u, \quad 0 < x < L \tag{9.1}$$

$$\alpha_1 u(0) - \beta_1 u'(0) = 0, \quad \alpha_2 u(L) + \beta_2 u'(L) = 0 \tag{9.2}$$

donde $\lambda$ es un parámetro, $\rho(x) > 0$ es una función peso y las relaciones (9.2) son condiciones en la frontera, en virtud de que se refieren a los puntos extremos $x = 0$ y $x = L$ del intervalo $[0, L]$. De acuerdo con las condiciones (8.85), se tiene que

$$p(x) \in C^1([0, L]), \quad q(x) \in C([0, L]), \quad p(x) > 0, \quad q(x) \geq 0$$

$$\alpha_1, \alpha_2 \geq 0, \quad \beta_1, \beta_2 \geq 0, \quad \alpha_1 + \beta_1 > 0, \quad \alpha_2 + \beta_2 > 0$$

Nótese que el dominio de definición $G_L$ del operador $\hat{L}$ está constituido por todas las funciones $u(x) \in \pounds_2([0, L], \rho)$ de la clase $C^2((0, L)) \cap C^1[(0, L)]$ que satisfacen las condiciones en la frontera (9.2) y $\hat{L}u \in \pounds_2([0, L], \rho)$.

Dado que el problema de Sturm-Liouville es un caso particular del problema de eigenvalores, ecuaciones (8.85) y (8.86), las propiedades de los eigenvalores y eigenfunciones siguen siendo válidas para el problema de Sturm-Liouville y, además, este problema tiene algunas propiedades específicas. Es claro que las propiedades de las eigenfunciones se extienden a las funciones especiales que surgen del problema. Dada su importancia, estas propiedades se listan aquí (véanse demostraciones, por ejemplo en [11] y [18]).

1. El conjunto de eigenvalores $\{\lambda_k\}$ del problema de Sturm-Liouville es contable y no tiene puntos límite finitos.

2. Los eigenvalores son reales, no negativos y simples.

3. Las eigenfunciones que corresponden a los eigenvalores diferentes son ortogonales entre sí sobre el intervalo $[0, L]$ con respecto a la función peso $\rho(x)$; las eigenfunciones pueden escogerse reales y ortonormales. Es importante notar que: a) si $p(0) = 0$, se puede eliminar la condición frontera en el punto $x = 0$; b) si $p(L) = 0$, se puede eliminar la condición frontera en el punto $x = L$; c) si $p(0) = p(L)$, se pueden reemplazar las condiciones (9.2) por $u(0) = u(L)$, $u'(0) = u'(L)$ sin alterar la ortogonalidad de las eigenfunciones.

4. Cada función del espacio lineal $G_L$ puede representarse en forma de serie de Fourier por eigenfunciones del operador $\hat{L}$; véase la ecuación (8.92).

## 9.2. El método de las series de potencias

En el capítulo 1 se vio que una ecuación diferencial ordinaria lineal homogénea, cuyos coeficientes son constantes, se puede resolver por medio de métodos algebraicos, y las soluciones son funciones elementales conocidas desde el Cálculo. Sin embargo, si los coeficientes son variables, la situación es más complicada y es posible que las soluciones no sean funciones elementales. La ecuación de Sturm-Liouville, ecuación (9.1), es diferencial lineal homogénea de segundo orden con coeficientes que, en general, no son constantes. Como las ecuaciones que generan funciones especiales, así como sus soluciones, desempeñan un papel importante en las Matemáticas aplicadas, en esta sección se considera un método para resolverlas. El método se conoce como **método de las series de potencias** o, en su forma generalizada, **método de Frobenius**.

La idea básica del método de las series de potencias para resolver ecuaciones diferenciales es sencilla y natural. Al darse una ecuación diferencial ordinaria, primero se representan todas las funciones dadas en esta ecuación en series de potencias de la variable independiente $x$. En muchos casos prácticos, las funciones dadas serán polinomios. Al suponer una solución en la forma de una serie de potencias

$$u(x) = \sum_{k=0}^{\infty} c_k x^k \tag{9.3}$$

se sustituye esta serie y las obtenidas al derivar término a término,

$$u'(x) = \sum_{k=1}^{\infty} k c_k x^{k-1}, \quad u''(x) = \sum_{k=2}^{\infty} k(k-1) c_k x^{k-2}, \quad \ldots \tag{9.4}$$

en la ecuación. Después de multiplicar las series término a término, se agrupan las potencias semejantes de $x$ y se iguala a cero la suma de los coeficientes de cada potencia de $x$ que se presente, empezando con los términos constantes. Esto da lugar a relaciones a partir de las cuales se pueden determinar sucesivamente los coeficientes desconocidos de la serie (9.3). El procedimiento se ilustra con un ejemplo simple.

**Ejemplo 9.1.** *Para resolver la ecuación*

$$u' - xu = 0$$

*suponemos una solución en la forma de la serie (9.3). Se sustituye (9.3) y la serie de la primera derivada (9.4) en la ecuación*

$$\sum_{k=0}^{\infty} (k+1) c_{k+1} x^k - \sum_{k=0}^{\infty} c_k x^{k+1} = 0$$

*donde se usó un desplazamiento simple del índice de la suma. Se agrupan las potencias semejantes de $x$, hallando*

$$c_1 + \sum_{k=1}^{\infty} \left((k+1) c_{k+1} - c_{k-1}\right) x^k = 0$$

*Al igualar a cero el coeficiente de cada potencia de $x$, se tiene*

$$c_{2m+1} = 0, \quad 2(m+1) c_{2(m+1)} - c_{2m} = 0 \quad \textit{para} \quad m = 0, 1, 2, \ldots$$

*Al escoger $c_0 = C$, donde $C$ es una constante arbitraria, para los coeficientes de las potencias par de $x$ se tiene*

$$c_{2(m+1)} = \frac{c_{2m}}{2(m+1)} = \frac{1}{2(m+1)} \frac{1}{2m} \frac{1}{2(m-1)} \cdots \frac{C}{2}$$

*Finalmente, la serie (9.3) queda*

$$u(x) = \sum_{k=0}^{\infty} c_{2k} x^{2k} = C \sum_{k=0}^{\infty} \frac{x^{2k}}{2^k k!} = C \exp\left(x^2/2\right)$$

*y se ve que se ha obtenido la solución general.*

Se ve que el método de las series de potencias proporciona soluciones en la forma de series y la aplicación del método requiere de operaciones como son la derivación, la adición y la multiplicación de series de potencias. Por tal razón

se recordarán los conceptos básicos del método y de dichas operaciones. Las tres operaciones son permisibles, en el sentido explicado en cada caso (demostraciones correspondientes se pueden ver, por ejemplo, en [12]). También se da una condición referente a la anulación de todos los coeficientes, la cual es un concepto básico en la aplicación del método de las series de potencias.

**Derivación término a término**. Sea

$$u(x) = \sum_{k=0}^{\infty} c_k x^k$$

una serie que converge para $|x| < R$, en donde el radio de convergencia $R > 0$. Entonces, la serie obtenida al derivar ésta término a término también converge para esas $x$ y la derivada es

$$u'(x) = \sum_{k=1}^{\infty} k c_k x^{k-1}$$

En otras palabras, una serie de potencias se puede derivar término a término.

**Adición término a término**. Sean

$$u(x) = \sum_{k=0}^{\infty} c_k x^k \quad \text{y} \quad v(x) = \sum_{k=0}^{\infty} a_k x^k$$

dos series con radios de convergencia $R$ y $\tilde{R}$, respectivamente. Suponemos que $R < \tilde{R}$. Entonces, la serie

$$\sum_{k=0}^{\infty} (c_k + a_k) x^k$$

converge y representa $u(x) + v(x)$ para cada $x$ del intervalo $|x| < R$. En otras palabras, dos series de potencias pueden sumarse término a término dentro del intervalo de convergencia menor.

**Multiplicación término a término**. Sean

$$u(x) = \sum_{k=0}^{\infty} c_k x^k \quad \text{y} \quad v(x) = \sum_{k=0}^{\infty} a_k x^k \tag{9.5}$$

dos series. Si se multiplica cada uno de los términos de la primera por cada uno de los términos de la segunda y se agrupan las potencias semejantes de $x$, se obtiene la serie

$$\begin{aligned} & a_0c_0 + (a_0c_1 + a_1c_0)\, x + (a_0c_2 + a_1c_1 + a_2c_0)\, x^2 + \cdots \\ = & \sum_{k=0}^{\infty} (a_0c_k + a_1c_{k-1} + \cdots + a_kc_0)\, x^k \end{aligned}$$

que se conoce como producto de Cauchy de las series (9.5). El producto de Cauchy de dos series de potencias (9.5) es absolutamente convergente para cada

$x$ dentro del intervalo de convergencia de cada una de éstas. Si las series tienen las sumas $u(x)$ y $v(x)$, respectivamente, el producto de Cauchy tiene la suma $s(x) = u(x)v(x)$.

**Anulación de todos los coeficientes**. Si una serie de potencias tiene un radio de convergencia positivo y su suma es igual a cero idénticamente en todo intervalo de convergencia, entonces cada uno de los coeficientes de la serie es cero.

Estas propiedades de las series de potencias constituyen la base teórica para la solución de ecuaciones diferenciales mediante el método de las series de potencias. Ahora, la cuestión que queda por averiguar es la existencia de soluciones en series de potencias para una ecuación diferencial dada. El criterio básico en términos de funciones analíticas (esto es, que se pueden expandir en serie de potencias alrededor de un punto $x = a$) se da por el siguiente teorema (sin la demostración).

**Teorema 9.1.** *Si las funciones $P(x)$, $Q(x)$ y $F(x)$ en la ecuación diferencial*

$$u'' + P(x)u' + Q(x)u = F(x) \tag{9.6}$$

*son analíticas en $x = a$, entonces toda solución $u(x)$ de la ecuación (9.6) es analítica en ese punto y, por consiguiente, se representa mediante una serie de potencias de $(x - a)$ con radio de convergencia $R > 0$.*

El teorema proporciona condiciones suficientes de la existencia de soluciones en series de potencias y, por tanto, pueden existir ecuaciones que no satisfacen las hipótesis del teorema; sin embargo, admiten soluciones de este tipo. Por ejemplo, la ecuación $xu'' + u' + xu = 0$ tiene una solución en serie de potencias de $x$, pero no satisface las hipótesis del teorema.

Se puede ver que el teorema se extiende a las ecuaciones diferenciales lineales de la forma $u^{(n)} + a_1(x)u^{(n-1)} + \cdots + a_n(x)u = r(x)$, en donde las funciones $a_i(x)$ $(i = 1, 2, \ldots, n)$ son analíticas en el punto $x = a$.

## 9.3. Puntos singulares

Resulta que en una gran variedad de problemas de importancia práctica, las ecuaciones diferenciales correspondientes tienen coeficientes que no son funciones analíticas. Dada la importancia práctica de las ecuaciones diferenciales de segundo orden, en esta sección se considera el concepto de punto singular de estas ecuaciones. Dicho concepto es de utilidad para clasificar las ecuaciones diferenciales e investigar la facilidad de una solución en forma de serie, que se considera en la siguiente sección.

Considérese la ecuación diferencial lineal homogénea de segundo orden

$$u'' + P(x)u' + Q(x)u = 0 \tag{9.7}$$

Si las funciones $P(x)$ y $Q(x)$ permanecen finitas en $x = a$, este punto es un punto ordinario. Cuando $P(x)$ y/o $Q(x)$ divergen a medida que $x \to a$, $a$ es un punto singular. Se distinguen dos clases de puntos singulares:

1. Si $P(x)$, $Q(x)$ o ambas divergen a medida que $x \to a$, pero $(x-a)P(x)$ y $(x-a)^2 Q(x)$ permanecen finitas en $x = a$, entonces $x = a$ se llama **punto singular regular** o **no esencial**.

2. Si $P(x)$, $Q(x)$ o ambas divergen de tal manera que $(x-a)P(x)$ y/o $(x-a)^2 Q(x)$ tienden a infinito a medida que $x \to a$, entonces $x = a$ se llama **punto singular irregular** o **esencial**.

Estas definiciones son consistentes para todos los valores finitos de $a$. Para analizar el punto $x = a \to \infty$, se hace el cambio de variable independiente $x = 1/z$ y, sustituyendo en la ecuación diferencial (9.7), se considera el límite $z \to 0$. Mediante el cambio de la variable en las derivadas, se tiene

$$\frac{du(x)}{dx} = \frac{du(z^{-1})}{dz}\frac{dz}{dx} = -\frac{1}{x^2}\frac{du(z^{-1})}{dz} = -z^2\frac{du(z^{-1})}{dz}$$

$$\begin{aligned} \frac{d^2u(x)}{dx^2} &= \frac{d}{dz}\left[\frac{du(x)}{dx}\right]\frac{dz}{dx} = -z^2\left[-2z\frac{du(z^{-1})}{dz} - z^2\frac{d^2u(z^{-1})}{dz^2}\right] \\ &= 2z^3\frac{du(z^{-1})}{dz} + z^4\frac{d^2u(z^{-1})}{dz^2} \end{aligned}$$

Utilizando estos resultados, la ecuación (9.7) se transforma en

$$z^4\frac{d^2u}{dz^2} + \left[2z^3 - z^2P(z^{-1})\right]\frac{du}{dz} + Q(z^{-1})\,u = 0 \tag{9.8}$$

Por tanto, los nuevos coeficientes

$$\tilde{P}(z) = \frac{2z^3 - z^2P(z^{-1})}{z^4} \quad \text{y} \quad \tilde{Q}(z) = \frac{Q(z^{-1})}{z^4} \tag{9.9}$$

determinan si el punto $x = \infty$ $(z = 0)$ es ordinario o singular a medida que $z \to 0$. Cuando las dos expresiones permanecen finitas, el punto $x = \infty$ es ordinario. Cuando $\tilde{P}(z)$ y/o $\tilde{Q}(z)$ divergen, pero $z\tilde{P}(z)$ y $z^2\tilde{Q}(z)$ permanecen finitas cuando $z \to 0$, el punto $x = \infty$ es singular regular. Si $\tilde{P}(z)$ y/o $\tilde{Q}(z)$ divergen con mayor rapidez que $1/z$ y/o $1/z^2$, respectivamente, el punto $x = \infty$ es singular esencial.

A continuación se listan algunas ecuaciones diferenciales de gran importancia práctica, con el fin de clasificar sus puntos singulares.

La **ecuación de Legendre**

$$(1-x^2)\,u'' - 2xu' + n(n+1)\,u = 0 \tag{9.10}$$

tiene

$$P(x) = -\frac{2x}{(1-x^2)}, \quad Q(x) = \frac{n(n+1)}{(1-x^2)}$$

Por tanto, los puntos $x = \pm 1$ son singulares regulares. Dado que

$$\tilde{P}(z) = \frac{2z^3 - z^2 P\left(z^{-1}\right)}{z^4} = \frac{2z^3 z^2}{z^4 \left(z^2 - 1\right)}, \quad \lim_{z \to 0} \tilde{P}(z) = 0$$

y

$$\tilde{Q}(z) = \frac{Q\left(z^{-1}\right)}{z^4} = \frac{n(n+1) z^2}{z^4 \left(z^2 - 1\right)}, \quad \lim_{z \to 0} \tilde{Q}(z) = \infty$$

pero

$$\lim_{z \to 0} z^2 \tilde{Q}(z) = -n(n+1)$$

el punto $x = \infty$ $(z = 0)$ es también un punto singular regular.

De manera similar se establece que la **ecuación de Chebyshev**

$$\left(1 - x^2\right) u'' - xu' + n^2 u = 0 \tag{9.11}$$

tiene tres puntos singulares regulares $x = \pm 1, x = \infty$.

La **ecuación hipergeométrica de Gauss**

$$x(1-x) u'' + [c - (1 + a + b) x] u' - abu = 0 \tag{9.12}$$

tiene tres puntos singulares regulares $x = 0, 1, \infty$. Como puede observarse, las tres ecuaciones, Legendre, Chebyshev e hipergeométrica, tienen tres puntos singulares regulares. Resulta que las soluciones de las ecuaciones de Legendre y Chebyshev (y muchas más funciones) se pueden expresar en términos de funciones hipergeométricas, que son soluciones de la hipergeométrica.

La **ecuación hipergeométrica confluente**

$$xu'' + (c - x) u' - au = 0 \tag{9.13}$$

que tiene una singularidad regular en $x = 0$ y una singularidad esencial en $x = \infty$, se puede obtener de la ecuación hipergeométrica de Gauss combinando dos de sus singularidades mediante un cambio de variable independiente. Haciendo el cambio $bx = z$, de la ecuación (9.12) se tiene

$$\frac{z}{b}\left(1 - \frac{z}{b}\right) b^2 \frac{d^2 u}{dz^2} + \left[c - (1 + a + b) \frac{z}{b}\right] b \frac{du}{dz} - abu = 0$$

Dividiendo entre $b$,

$$z\left(1 - \frac{z}{b}\right) \frac{d^2 u}{dz^2} + \left[c - z + \frac{1 + a}{b} z\right] \frac{du}{dz} - au = 0$$

Tendiendo $b \to \infty$, la última se transforma en la ecuación hipergeométrica confluente

$$z \frac{d^2 u}{dz^2} + (c - z) \frac{du}{dz} - au = 0$$

Dicho cambio de variable independiente tuvo el efecto de combinar (confluir) las dos singularidades regulares en $x = 0$ y $x = 1$ en una singularidad regular en $z = 0$ y convertir la singularidad en $x = \infty$ de regular a esencial.

La **ecuación de Hermite**

$$u'' - 2xu' + \lambda u = 0 \tag{9.14}$$

tiene una singularidad esencial en $x = \infty$.

Un **oscilador armónico simple**

$$u'' + \omega^2 u = 0 \tag{9.15}$$

tiene una singularidad esencial en $x = \infty$.

La **ecuación de Laguerre**

$$xu'' + (1 - x)\, u' + \lambda u = 0 \tag{9.16}$$

tiene una singularidad regular en $x = 0$ y una esencial en $x = \infty$.

La **ecuación de Bessel**

$$x^2 u'' + xu' + \left(x^2 - \nu^2\right) u = 0 \tag{9.17}$$

tiene una singularidad regular en $x = 0$ y otra esencial en $x = \infty$.

**Problema.** *Demuestre la existencia de las singularidades enunciadas de las ecuaciones anteriores.*

## 9.4. El método de Frobenius

En esta sección se desarrolla un **método de solución en series** de la ecuación diferencial lineal homogénea de segundo orden. El método siempre resulta adecuado considerando que el punto de expansión no es más que un punto singular regular de la ecuación diferencial. En la Física e Ingeniería, esta condición bastante moderada casi siempre se satisface.

De la teoría de ecuaciones diferenciales se conoce que, si $u_1(x)$ y $u_2(x)$ son dos soluciones linealmente independientes de una ecuación diferencial lineal homogénea de segundo orden, ecuación (9.7), la solución general de dicha ecuación es entonces

$$u(x) = C_1 u_1(x) + C_2 u_2(x)$$

La solución general de la ecuación no homogénea correspondiente, ecuación (9.6), es

$$u(x) = C_1 u_1(x) + C_2 u_2(x) + u_p(x)$$

en donde $u_p(x)$ es una solución particular cualquiera de (9.6). Admitiendo que una solución particular se puede obtener con la técnica de la transformada de Laplace o con la de la función de Green, desarrollaremos en esta sección un método de series de potencias para obtener las soluciones fundamentales $u_1(x)$ y $u_2(x)$.

Resulta que varias ecuaciones diferenciales de segundo orden, de gran importancia en muchas aplicaciones, tienen coeficientes que no son analíticas en $x = 0$, pero son tales que se puede aplicar el siguiente teorema (presentado sin demostración).

**Teorema 9.2.** *Cualquier ecuación diferencial de la forma*

$$u'' + \frac{a(x)}{x}u' + \frac{b(x)}{x^2}u = 0 \tag{9.18}$$

*en donde las funciones $a(x)$ y $b(x)$ son analíticas en $x = 0$, tiene al menos una solución que se puede representar en la forma de serie*

$$u(x) = x^r \sum_{k=0}^{\infty} c_k x^k, \quad c_0 \neq 0 \tag{9.19}$$

*en la que el exponente $r$ puede ser cualquier número (real o complejo) y que se elige de modo que $c_0 \neq 0$.*

El punto clave es que en la ecuación (9.19) se tenga una serie de potencias multiplicada por una sola potencia de $x$ cuyo exponente no esté restringido a ser un entero no negativo. Al método para resolver la ecuación (9.18) en la forma (9.19) se le llama **método de Frobenius**.

Para resolver la ecuación (9.18), la escribimos en una forma un tanto más conveniente

$$x^2u'' + xa(x)u' + b(x)u = 0 \tag{9.20}$$

Primero se desarrollan $a(x)$ y $b(x)$ en series de potencias

$$a(x) = \sum_{k=0}^{\infty} a_k x^k, \quad b(x) = \sum_{k=0}^{\infty} b_k x^k$$

Después se deriva (9.19) término a término, hallando

$$\begin{aligned} u'(x) &= x^{r-1} \sum_{k=0}^{\infty} (r+k) c_k x^k \\ u''(x) &= x^{r-2} \sum_{k=0}^{\infty} (r+k)(r+k-1) c_k x^k \end{aligned}$$

Al sustituir todas estas series en la ecuación (9.20), se obtiene

$$\begin{aligned} & x^r \sum_{k=0}^{\infty} (r+k)(r+k-1) c_k x^k + \\ & x^r \left( \sum_{m=0}^{\infty} a_m x^m \right) \left( \sum_{k=0}^{\infty} (r+k) c_k x^k \right) + \\ & x^r \left( \sum_{m=0}^{\infty} b_m x^m \right) \left( \sum_{k=0}^{\infty} c_k x^k \right) = 0 \end{aligned}$$

Ahora se igualan a cero las sumas de los coeficientes de cada potencia de $x$, como antes. Esto da lugar a un sistema de ecuaciones, llamado relaciones de

recurrencia, que contienen a los coeficientes desconocidos $c_k$. La menor potencia es $x^r$ y la ecuación correspondiente es

$$[r(r-1)+a_0r+b_0]\,c_0=0$$

De donde, en virtud de la suposición $c_0 \neq 0$, debe cumplirse la igualdad

$$r^2+(a_0-1)\,r+b_0=0 \tag{9.21}$$

Esta importante ecuación cuadrática se llama **ecuación indicial** de la ecuación diferencial (9.18).

El método que se está describiendo permite obtener una base de soluciones. Una de las dos funciones base siempre será de la forma (9.19), en donde $r$ es una raíz de la ecuación (9.21). La forma de la otra se determina por la ecuación indicial. Dependiendo de las raíces se distinguen tres casos, los cuales se analizarán por separado:

1. Raíces distintas $r_1$ y $r_2$ que no difieren en un entero. En este caso, las dos soluciones serán de la forma (9.19) con $r=r_1, r_2$ y coeficientes diferentes.

2. Raíz doble. En este caso, la primera solución será de la forma (9.19), en tanto que la segunda debe contener un término logarítmico.

3. Raíces distintas $r_1$ y $r_2$ que difieren en un entero. En este caso, la primera solución será de la forma (9.19), en tanto que la segunda puede contener o no un término logarítmico.

**Ejemplo 9.2.** *No es de sorprender que el caso de una raíz doble cree una situación especial al determinar dos soluciones linealmente independientes. Una ilustración sencilla de este hecho se da por la ecuación de Cauchy-Euler*

$$x^2u''+axu'+bu=0$$

*en donde $a$ y $b$ son constantes. Es fácil adivinar que esta ecuación se puede resolver por la sustitución $u=x^r$. Esto conduce a la ecuación auxiliar*

$$r^2+(a-1)\,r+b=0$$

*que es la ecuación indicial. Si las raíces son $r_1$ y $r_2 \neq r_1$, una base de soluciones está dada por*

$$u_1=x^{r_1}, \quad u_2=x^{r_2}$$

*Esto corresponde a los casos 1 o 3. Si $r_1=r_2=r$, una base es*

$$u_1=x^r, \quad u_2=x^r\ln x$$

*que corresponde al caso 2.*

**Caso 1. Raíces distintas que no difieren en un entero**

Éste es el caso más sencillo. Sean $r_1$ y $r_2$ las raíces de (9.21). Sustituyendo $r = r_1$ en el sistema de ecuaciones que contienen a los coeficientes desconocidos $c_k$ $(k = 0, 1, 2, \ldots)$ y resolviendo dichas ecuaciones sucesivamente, se obtiene una solución

$$u_1(x) = x^{r_1} \sum_{k=0}^{\infty} c_k x^k \tag{9.22}$$

Otra solución

$$u_2(x) = x^{r_2} \sum_{k=0}^{\infty} \tilde{c}_k x^k \tag{9.23}$$

se obtiene al sustituir $r = r_2$ en las relaciones de recurrencia y al determinar sucesivamente los coeficientes $\tilde{c}_k$ $(k = 0, 1, 2, \ldots)$. Se puede demostrar la independencia lineal de $u_1(x)$ y $u_2(x)$ a partir del hecho de que la razón $u_1(x)/u_2(x)$ no es constante porque $r_1 - r_2$ no es un entero. Por tanto, $u_1(x)$ y $u_2(x)$ forman una base de (9.18) y (9.20) en un intervalo abierto en el que las dos series converjan. La solución general, por supuesto, es $u(x) = C_1 u_1(x) + C_2 u_2(x)$, en donde $C_1$ y $C_2$ son constantes arbitrarias.

**Ejercicio.** *Resolver la ecuación* $2xu'' + (1 + x)u' + u = 0$ *por el método de Frobenius.*

**Caso 2. Raíz doble**

La ecuación indicial (9.21) tiene una raíz doble si y sólo si $(a_0 - 1)^2 - 4b_0 = 0$. Por tanto, $r = (1 - a_0)/2$. En este caso se puede determinar una primera solución

$$u_1(x) = x^r \sum_{k=0}^{\infty} c_k x^k, \quad r = (1 - a_0)/2 \tag{9.24}$$

como antes.

Una segunda solución linealmente independiente es

$$u_2(x) = u_1(x) \ln x + x^r \sum_{k=1}^{\infty} A_k x^k, \quad x > 0 \tag{9.25}$$

Para demostrar el hecho, se busca una segunda solución de la forma

$$u_2(x) = v(x)\, u_1(x) \tag{9.26}$$

en donde $v(x)$ es una función desconocida. Introduciendo (9.26) y sus derivadas

$$u_2'(x) = v'u_1 + vu_1', \quad u_2''(x) = v''u_1 + 2v'u_1' + vu_1''$$

en la ecuación (9.20), se obtiene

$$x^2(v''u_1 + 2v'u_1' + vu_1'') + xa(x)(v'u_1 + vu_1') + b(x)\,vu_1 = 0$$

Como $u_1$ es una solución de (9.20), la ecuación anterior se reduce a

$$x^2 u_1 v'' + 2x^2 u_1' v' + xa(x) u_1 v' = 0$$

o

$$v'' + \tilde{b}(x) v' = 0 \tag{9.27}$$

donde

$$\tilde{b}(x) \equiv 2\frac{u_1'}{u_1} + \frac{a(x)}{x} \tag{9.28}$$

La solución general de la ecuación (9.27) es

$$v(x) = C_1 \int \exp\left(-\int \tilde{b}(x)\, dx\right) dx + C_2 \tag{9.29}$$

donde $C_1$ y $C_2$ son dos constantes arbitrarias. En lo que sigue, escogemos una solución particular con $C_1 = 1$ y $C_2 = 0$. Sustituyendo las series de potencias de $a(x)$, $u_1(x)$ y $u_1'(x)$ en $\tilde{b}(x)$ y usando la igualdad $r = (1 - a_0)/2$, se tiene

$$\tilde{b}(x) = \frac{1}{x} + \cdots$$

donde los puntos suspensivos designan términos que son constantes o que contienen potencias positivas de $x$. Después, efectuando la integración en (9.29), finalmente se obtiene una solución particular de la ecuación (9.27) de la forma

$$v(x) = \int \exp\left(-\int \tilde{b}(x)\, dx\right) dx = \ln x + \sum_{k=1}^{\infty} \tilde{A}_k x^k \tag{9.30}$$

La última función sustituida en (9.26) conduce a la fórmula (9.25). Nótese que el término que contiene $\ln x$ siempre está presente en una segunda solución.

**Ejercicio.** *Resuelva la ecuación* $x(x-1)u'' + (3x-1)u' + u = 0$.

### Caso 3. Raíces que difieren en un entero

Si las raíces $r_1$ y $r_2$ de la ecuación indicial difieren en un entero, digamos $r_1 - r_2 = p$, en donde $p$ es un entero positivo, siempre es posible determinar una solución como antes

$$u_1(x) = x^{r_1} \sum_{k=0}^{\infty} c_k x^k \tag{9.31}$$

correspondiente a la raíz $r_1$.

Una segunda solución linealmente independiente es

$$u_2(x) = K_p u_1(x) \ln x + x^{r_2} \sum_{k=0}^{\infty} A_k x^k, \quad x > 0 \tag{9.32}$$

Aquí es posible que $K_p \neq 0$ o $K_p = 0$, de modo que se puede tener o no el término con $\ln x$, dependiendo de la ecuación diferencial dada. Los primeros

pasos de demostración son idénticos a los del caso 2. La distinción aparece al sustituir las series de potencias de $a(x)$, $u_1(x)$ y $u_1'(x)$ en $\tilde{b}(x)$, resultando en

$$\tilde{b}(x) = \left( \frac{2r_1 + a_0}{x} + \cdots \right)$$

Aquí, y en lo que sigue, los puntos suspensivos designan términos que son constantes o que contienen potencias positivas de $x$. Del Álgebra elemental se sabe que el coeficiente $(a_0 - 1)$ en la ecuación (9.21) es igual a $-(r_1 + r_2)$. Por lo tanto, usando la igualdad $r_1 - r_2 = p$, se tiene $2r_1 + a_0 = 1 + p$ y

$$\tilde{b}(x) = \left( \frac{1+p}{x} + \cdots \right)$$

Integrando, se tiene

$$\begin{aligned}
\exp\left( -\int \tilde{b}(x)\, dx \right) &= C_1 \left( \frac{1}{x} \right)^{1+p} \left( \sum_{k=0}^{\infty} \tilde{A}_k x^k \right) \\
\int \exp\left( -\int \tilde{b}(x)\, dx \right) dx &= C_1 \int \left( \frac{1}{x} \right)^{1+p} \left( \sum_{k=0}^{\infty} \tilde{A}_k x^k \right) dx \\
&= C_1 \left( -\frac{\tilde{A}_0}{p x^p} - \cdots + \tilde{A}_p \ln x + \tilde{A}_{p+1} x + \cdots \right)
\end{aligned}$$

De donde una solución particular con $C_1 = 1$ y $C_2 = 0$ es

$$\begin{aligned}
v(x) &= \int \exp\left( -\int \tilde{b}(x)\, dx \right) dx \\
&= C_1 \int \left( \frac{1}{x} \right)^{1+p} \left( \sum_{k=0}^{\infty} \tilde{A}_k x^k \right) dx \\
&= C_1 \left( -\frac{\tilde{A}_0}{p x^p} - \cdots + \tilde{A}_p \ln x + \tilde{A}_{p+1} x + \cdots \right)
\end{aligned}$$

Se multiplica esta expresión por la serie (9.31) y, dado que $r_1 = r_2 + p$, se obtiene (9.32), en donde $K_p = C_1 \tilde{A}_p$. Cuando el coeficiente $\tilde{A}_p = 0$, el término con $\ln x$ no aparece en (9.32). Esto completa la demostración en el caso 3.

**Problema.** *Resuelva la ecuación* $x^2 (x^2 - 1) u'' - x (x^2 + 1) u' + (x^2 + 1) u = 0$, *demostrando que la segunda solución contiene el término logarítmico.*

**Problema.** *Resuelva la ecuación* $x^2 u'' + x u' + (x^2 - 1/4) u = 0$, *demostrando que la segunda solución no tiene el término logarítmico.*

## 9.5. Funciones y polinomios especiales

### 9.5.1. Polinomios y funciones asociadas de Legendre

#### Funciones de Legendre

La ecuación diferencial de Legendre

$$\left(1-x^2\right)u''-2xu'+n\left(n+1\right)u=0, \quad -1\leq x\leq 1 \tag{9.33}$$

surge en numerosos problemas físicos con valores en la frontera, en particular para esferas. El parámetro $n$ de esta ecuación es un número real dado. Cualquier solución de (9.33) se llama **función de Legendre**.

Nótese que la ecuación (9.33), escrita en la forma

$$\left[\left(1-x^2\right)u'\right]'+\lambda u=0 \tag{9.34}$$

representa un problema de eigenvalores para el operador diferencial $\hat{L}$

$$\hat{L}u\equiv-\left(p\,u'\right)'+qu=\lambda\rho u$$

donde el parámetro $\lambda=n\left(n+1\right)$ es un eigenvalor del problema y

$$p\left(x\right)=1-x^2, \quad \rho\left(x\right)=1, \quad q\left(x\right)=0, \quad -1=a\leq x\leq b=1$$

Dado que la función $p\left(x\right)=1-x^2$ se anula en los extremos del intervalo de definición de soluciones, $p\left(\pm1\right)=0$, la ortogonalidad de funciones de Legendre no depende de las condiciones en la frontera (véase el problema de Sturm-Liouville, sección 9.1).

Dividiendo la ecuación (9.33) entre $\left(1-x^2\right)$, se obtiene la forma estándar (9.6), y se ve que los coeficientes de la ecuación resultante son analíticos en $x=0$, de modo que es posible aplicar el método de las series de potencias. Al sustituir

$$u\left(x\right)=\sum_{k=0}^{\infty}c_kx^k \tag{9.35}$$

y sus derivadas en la ecuación (9.33), se obtiene

$$\left(1-x^2\right)\sum_{k=2}^{\infty}k\left(k-1\right)c_kx^{k-2}-2x\sum_{k=1}^{\infty}kc_kx^{k-1}+\lambda\sum_{k=0}^{\infty}c_kx^k=0$$

donde la constante $\lambda=n\left(n+1\right)$. Si (9.35) tiene que ser una solución de (9.33), la suma de los coeficientes de cada potencia de $x$ debe ser cero. Esto da la relación de recurrencia

$$\left(k+2\right)\left(k+1\right)c_{k+2}+\left[\lambda-k\left(k+1\right)\right]c_k=0, \quad k=0,1,2,\ldots \tag{9.36}$$

o bien,

$$c_{k+2}=\frac{k\left(k+1\right)-\lambda}{\left(k+2\right)\left(k+1\right)}c_k, \quad k=0,1,2,\ldots \tag{9.37}$$

Dicha relación no determina los coeficientes $c_0$ y $c_1$, que se dejan como constantes arbitrarias. De la ecuación (9.37) se ve que los coeficientes $c_k$ se separan en dos grupos, los coeficientes de índice par e impar, cada grupo representando una solución de (9.33). Introduciendo valores de los coeficientes en (9.35), se obtiene

$$u(x) = c_0 u_1(x) + c_1 u_{2(x)} \tag{9.38}$$

en donde

$$\begin{aligned} u_1(x) &= c_0^{-1} \sum_{k=0}^{\infty} c_{2k} x^{2k} \qquad (9.39) \\ &= 1 - \frac{\lambda}{2!} x^2 - \frac{\lambda(2 \cdot 3 - \lambda)}{4!} x^4 - \cdots \end{aligned}$$

y

$$\begin{aligned} u_2(x) &= c_1^{-1} \sum_{k=0}^{\infty} c_{2k+1} x^{2k+1} \qquad (9.40) \\ &= x + \frac{1 \cdot 2 - \lambda}{3!} x^3 + \frac{(1 \cdot 2 - \lambda)(3 \cdot 4 - \lambda)}{5!} x^5 + \cdots \end{aligned}$$

Estas series convergen para $|x| < 0$. Es fácil ver que la función $u_1(x)$ es par, en tanto que $u_2(x)$ es impar y, por tanto, son soluciones linealmente independientes. De donde, la función (9.38) es una solución general de la ecuación de Legendre (9.33) en el intervalo $-1 < x < 1$.

## Polinomios de Legendre

De la ecuación (9.37) se ve que, cuando el parámetro $\lambda$ es igual a $\lambda = n(n+1)$, en donde $n$ es un número natural, todos los coeficientes $c_{n+2s} = 0$ $(s = 1, 2, \ldots)$. De donde, si $n$ es par, $u_1(x)$ se reduce a un polinomio de grado $n$, quedando $u_2(x)$ una serie. Si $n$ es impar, $u_2(x)$ se reduce a un polinomio de grado $n$, quedando $u_1(x)$ una serie. Dichos polinomios, multiplicados por algunas constantes, se llaman **polinomios de Legendre** de grado $n$. Dada su gran importancia práctica, estos polinomios se considerarán con más detalle.

Con este fin, resolviendo (9.37) para $c_k$, se obtiene

$$c_k = \frac{(k+2)(k+1)}{k(k+1) - n(n+1)} c_{k+2}, \quad k \leq n-2$$

Entonces, se pueden expresar todos los coeficientes que no se anulan en términos del coeficiente $c_n$ de la mayor potencia de $x$ del polinomio. En principio, el coeficiente $c_n$ es arbitrario. Por razones de normalización de los polinomios de Legendre, se acostumbra elegir

$$c_n = \frac{(2n)!}{2^n (n!)^2} = \frac{1 \cdot 3 \cdots (2n-1)}{n!}, \quad n = 0, 1, 2, \ldots$$

En este caso, todos los polinomios tendrán el valor 1 cuando $x = 1$. Las soluciones parciales resultantes de la ecuación diferencial de Legendre (9.33) se llaman polinomios de Legendre de grado $n$ y se denotan por $P_n(x)$ $(n = 0, 1, 2, \ldots)$. En particular,

$$P_0(x) = 1, \qquad P_1(x) = x,$$
$$P_2(x) = \left(3x^2 - 1\right)/2, \quad P_3(x) = \left(5x^3 - 3x\right)/2$$

La siguiente expresión, conocida como **fórmula de Rodrigues**, permite generar los polinomios de Legendre

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} \left[ \left(x^2 - 1\right)^n \right], \quad -1 \leq x \leq 1 \tag{9.41}$$

Para comprobarla, definimos la función $w_n(x) = \left(x^2 - 1\right)^n$. Al diferenciar la identidad

$$\left(x^2 - 1\right) w_n' - 2nxw_n = 0$$

$(n+1)$ veces, se obtiene

$$\left(x^2 - 1\right) w_n^{(n+2)} + 2xw_n^{(n+1)} - n(n+1) w_n^{(n)} = 0$$

Entonces, la función $w_n^{(n)}(x)$ y, por tanto, el polinomio $P_n(x)$ representado por la fórmula de Rodrigues, es una solución de (9.33), en donde el parámetro $n$ es un número natural.

La representación integral de los polinomios de Legendre está dada por la siguiente fórmula (sin demostración):

$$P_n(\cos\theta) = \frac{1}{\pi} \int_0^\pi (\cos\theta \pm i \operatorname{sen}\theta \cos\phi)^n \, d\phi = \frac{2}{\pi} \int_0^\pi \frac{\cos\left(n + \frac{1}{2}\right)\phi d\phi}{\sqrt{2(\cos\phi - \cos\theta)}} \tag{9.42}$$

en donde la variable independiente $x = \cos\theta$.

La función

$$G(x, h) = \frac{1}{\sqrt{1 - 2hx + h^2}} \tag{9.43}$$

se conoce como **función generadora** de los polinomios de Legendre, en virtud de que su desarrollo en serie de potencias de $h$ es

$$G(x, h) = \frac{1}{\sqrt{1 - 2hx + h^2}} = \sum_{n=0}^{\infty} P_n(x) h^n, \qquad |h| < 1 \tag{9.44}$$

Para verificar la última relación podemos seguir los siguientes pasos. Se sustituye $y = 2hx - h^2$ en la expansión de $(1 - y)^{-1/2}$ en serie de potencias de $y$. Después, desarrollando las potencias de $\left(2hx - h^2\right)$ y agrupando todos los términos que contengan $h^n$, se obtiene que la suma de estos términos es $P_n(x) h^n$.

La función generadora, ecuación (9.43), satisface la relación

$$G(x, h) = G(-x, -h)$$

es decir,

$$\sum_{n=0}^{\infty} h^n P_n(x) = \sum_{n=0}^{\infty} (-h)^n P_n(-x)$$

A partir de esta expresión se obtiene la paridad de los polinomios de Legendre

$$P_n(-x) = (-1)^n P_n(x) \tag{9.45}$$

A partir de la función generadora $G(x,h)$ se obtienen **relaciones de recurrencia**:

$$(n+1) P_{n+1} - (2n+1) x P_n + n P_{n-1} = 0 \tag{9.46}$$

$$(2n+1) P_n = P'_{n+1} - P'_{n-1} \tag{9.47}$$

$$(x^2 - 1) P_n = nxP_n - nP_{n-1} \tag{9.48}$$

$$= (n+1) P_{n+1} - (n+1) x P_n \tag{9.49}$$

Derivando parcialmente la igualdad (9.44) con respecto a $h$ y $x$, y multiplicando después por $(1 - 2hx + h^2)$, obtenemos las igualdades

$$(x-h) \sum_{n=0}^{\infty} P_n(x) h^n = (1 - 2hx + h^2) \sum_{n=0}^{\infty} n P_n(x) h^{n-1} \tag{9.50}$$

$$h \sum_{n=0}^{\infty} P_n(x) h^n = (1 - 2hx + h^2) \sum_{n=0}^{\infty} P'_n(x) h^n \tag{9.51}$$

Comparando los coeficientes de las mismas potencias de $h$ en la ecuación (9.50), se tiene

$$xP_n - P_{n-1} = (n+1) P_{n+1} - 2xnP_n + (n-1) P_{n-1}$$

de donde sigue la relación (9.46). De modo similar, de la ecuación (9.51) se tiene

$$P_n = P'_{n+1} - 2xP'_n + P'_{n-1}$$

Derivando la igualdad (9.50), se tiene

$$(n+1) P'_{n+1} - (2n+1) P_n - (2n+1) x P'_n + n P'_{n-1} = 0$$

Luego, excluyendo el término $xP'_n$ de las dos últimas ecuaciones, se obtiene la relación (9.47).

**Problema.** *Demuestre las relaciones de recurrencia, ecuaciones (9.48) y (9.49).*

Los polinomios de Legendre, por el hecho de ser eigenfunciones correspondientes a diferentes eigenvalores $\lambda_n = n(n+1)$ del problema (9.34), son ortogonales entre sí:

$$\int_{-1}^{1} P_n(x) P_m(x) \, dx = 0 \quad \text{si} \quad n \neq m \tag{9.52}$$

Usando la relación de recurrencia (9.46) y la ortogonalidad de los polinomios de Legendre, se demuestra que

$$\int_{-1}^{1} P_n^2(x)\,dx = \frac{2}{2n+1} \tag{9.53}$$

Por lo tanto, las funciones

$$\Phi_n(x) = \sqrt{n+\frac{1}{2}}\,P_n(x) \tag{9.54}$$

son ortonormales, es decir,

$$\int_{-1}^{1} \Phi_n(x)\,\Phi_m(x)\,dx = \delta_{nm} \tag{9.55}$$

El conjunto de polinomios de Legendre es completo en el espacio de funciones $\pounds_2(-1,1)$, es decir, cada función $f(x) \in \pounds_2(-1,1)$ puede representarse mediante la serie de Fourier de polinomios de Legendre

$$f(x) = \sum_{n=0}^{\infty} \frac{2n+1}{2}(f, P_n)\,P_n(x) = \sum_{n=0}^{\infty}(f, \Phi_n)\,\Phi_n(x) \tag{9.56}$$

que es convergente en $\pounds_2(-1,1)$.

**Funciones asociadas de Legendre**

La ecuación diferencial

$$\left[\left(1-x^2\right)y'\right]' - \frac{m^2}{1-x^2}y + n(n+1)\,y = 0 \tag{9.57}$$

$$p(x) = 1-x^2, \quad \rho(x) = 1, \quad q(x) = \frac{m^2}{1-x^2}, \quad -1 = a \leq x \leq b = 1$$

es una ecuación asociada de Legendre, en donde $m$ es un número natural dado $0 \leq m \leq n$. Las soluciones de esta ecuación se llaman **funciones asociadas de Legendre**. Haciendo la sustitución

$$y(x) = \left(1-x^2\right)^{m/2} u(x)$$

se tiene la ecuación

$$\left(1-x^2\right)u'' - 2x\,(m+1)\,u' + \left[\lambda - m\,(m+1)\right]u = 0$$

que se satisface por

$$u = \frac{d^m}{dx^m} P_n(x)$$

Lo último se puede demostrar, derivando la ecuación de Legendre, ecuación (9.33), $m$ veces. Entonces, la función

$$P_n^m(x) = \left(1-x^2\right)^{m/2} \frac{d^m}{dx^m} P_n(x), \quad 0 \leq m \leq n \tag{9.58}$$

conocida como **función asociada de Legendre**, es una solución de (9.57).

Para cada número natural $m$, las funciones asociadas de Legendre son ortogonales

$$\int_{-1}^{1} P_n^m(x) P_k^m(x)\, dx = \begin{cases} 0 & k \neq n \\ \frac{2}{2n+1} \frac{(n+m)!}{(n-m)!} & k = n \end{cases} \tag{9.59}$$

## 9.5.2. Polinomios de Hermite

En mecánica cuántica es muy conocido el conjunto de polinomios de Hermite que surgen como soluciones del problema de eigenvalores de un oscilador armónico cuántico. El problema de Sturm-Liouville correspondiente se describe por la ecuación

$$\left(e^{-x^2} u'\right)' + \lambda e^{-x^2} u = 0 \tag{9.60}$$

donde el operador diferencial $\hat{L}u = -\left(pu'\right)' + qu$ y la función peso $\rho(x)$ se caracterizan mediante

$$p(x) = e^{-x^2}, \quad \rho(x) = e^{-x^2}, \quad q(x) = 0; \quad -\infty = a < x < b = \infty \tag{9.61}$$

De (9.60) se obtiene la ecuación

$$u'' - 2xu' + \lambda u = 0 \tag{9.62}$$

Nótese que ésta no es la forma canónica del problema de Sturm-Liouville. La última ecuación es de la forma de la ecuación (9.6) con coeficientes analíticos en $x = 0$. Por tanto, toda solución $u(x)$ de la ecuación (9.62) es analítica en ese punto y, por consiguiente, se representa mediante una serie de potencias de $x$ con radio de convergencia $R > 0$. Aplicando el método de las series de potencias para la solución de la ecuación (9.62), para los eigenvalores $\lambda_n = 2n$ $(n = 0, 1, 2, \ldots)$, se obtienen las soluciones (eigenfunciones del problema (9.60)) en forma polinomial

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n}\left(e^{-x^2}\right), \quad n = 0, 1, 2, \ldots \tag{9.63}$$

Estas funciones se llaman **polinomios de Hermite**.

Como sucede con muchas funciones especiales, la literatura correspondiente presenta más de una notación, y a veces se definen como polinomios de Hermite las funciones

$$He_n(x) = (-1)^n e^{x^2/2} \frac{d^n}{dx^n}\left(e^{-x^2/2}\right), \quad n = 0, 1, 2, \ldots$$

Estos polinomios son soluciones de la ecuación

$$u'' - xu' + nu = 0, \quad n = 0, 1, 2, \ldots$$

Los polinomios de Hermite correspondientes a eigenvalores diferentes son ortogonales. El sistema completo de funciones ortonormales es

$$\Phi_n(x) = e^{\frac{-x^2}{2}} H_n(x) \frac{1}{2^{\frac{n}{2}} \pi^{\frac{1}{4}} (n!)^{\frac{1}{2}}}, \quad n = 0, 1, 2, \ldots \tag{9.64}$$

$$(\Phi_n, \Phi_m) = \delta_{nm}, \quad n, m = 0, 1, 2, \ldots \tag{9.65}$$

La función generadora de los polinomios de Hermite es la función exponencial

$$G(x, h) = \exp\left(-h^2 + 2xh\right) = \sum_{n=0}^{\infty} \frac{H_n(x) h^n}{n!} \tag{9.66}$$

La función generadora es par, $G(x, h) = G(-x, -h)$, es decir,

$$\sum_{n=0}^{\infty} \frac{H_n(x) h^n}{n!} = \sum_{n=0}^{\infty} \frac{H_n(-x)(-1)^n h^n}{n!}$$

Entonces,

$$H_n(-x) = (-1)^n H_n(x) \tag{9.67}$$

Utilizando la función generadora se pueden comprobar varias relaciones de recurrencia para los polinomios de Hermite. Por ejemplo, derivando la ecuación (9.66) se obtiene

$$H_n' = 2nH_{n-1} \tag{9.68}$$

Además,

$$H_{n+1} = 2xH_n - 2nH_{n-1} \tag{9.69}$$

## 9.5.3. Polinomios de Laguerre y polinomios asociados

### Polinomios de Laguerre

Los polinomios y funciones asociadas de Laguerre están estrechamente relacionados con el problema cuántico de un átomo de hidrógeno. El problema de Sturm-Liouville que da origen a los polinomios de Laguerre se describe por la siguiente ecuación

$$\left(xe^{-x}u'\right)' + \lambda e^{-x}u = 0 \tag{9.70}$$

donde el operador diferencial $\hat{L}u = -(pu')' + qu$ y la función peso $\rho(x)$ se caracterizan mediante

$$p(x) = xe^{-x}, \quad \rho(x) = e^{-x}, \quad q(x) = 0, \quad 0 = a \leq x < b = \infty \tag{9.71}$$

De (9.70) se obtiene la ecuación

$$xu'' + (1-x)\,u' + \lambda u = 0 \tag{9.72}$$

que se puede resolver por el método de Frobenius (véase la ecuación (9.18)). La solución de la ecuación (9.72) se busca en forma de serie (9.19). En este caso, la ecuación indicial

$$r^2 = 0$$

tiene la raíz doble $r_{1,2} = 0$. Por tanto, la primera solución de (9.72) es de la forma (9.19). Para eigenvalores $\lambda_n = n$ $(n = 0, 1, 2, \ldots)$, dicha solución se reduce a los polinomios

$$\begin{aligned} L_n(x) &= \sum_{k=0}^{n} \frac{(-1)^k}{k!} \binom{n}{k} x^k \\ &= 1 - nx + \frac{n(n-1)}{4} x^2 - \cdots + \frac{(-1)^n}{n!} x^n \end{aligned} \tag{9.73}$$

llamados **polinomios de Laguerre** de grado $n$. Estos polinomios tienen las siguientes representaciones: la diferencial

$$n! L_n(x) = e^x \frac{d^n}{dx^n} \left(e^{-x} x^n\right) \tag{9.74}$$

y la integral

$$n! L_n(x) = \int_0^\infty e^{x-t} t^n J_0\left(2\sqrt{xt}\right) dt \tag{9.75}$$

donde $J_0$ es la función de Bessel de primera clase de orden cero (véase la sección 9.4.5).

El conjunto de todos los polinomios de Laguerre es ortogonal sobre el intervalo $0 \leq x < \infty$, con respecto a la función peso $\rho(x) = e^{-x}$, es decir,

$$(L_n, L_m)_\rho = \int_0^\infty L_n(x)\, L_m(x)\, \rho(x)\, dx = \delta_{nm}, \quad n, m = 0, 1, 2, \ldots \tag{9.76}$$

Las funciones ortonormales del problema (9.70) son

$$\Phi_n(x) = e^{-x/2} L_n(x), \quad (\Phi_n, \Phi_m) = \delta_{nm}, \quad n, m = 0, 1, 2, \ldots \tag{9.77}$$

La función generadora de los polinomios de Laguerre es

$$\frac{1}{1-t} \exp\left(\frac{-xt}{1-t}\right) = \sum_{n=0}^{\infty} L_n(x)\, t^n, \quad |t| < 1 \tag{9.78}$$

Mediante la diferenciación de la función generadora, ecuación (9.78), con respecto a $x$ y $t$, se demuestran las siguientes relaciones de recurrencia:

$$(n+1)\, L_{n+1} - (2n + 1 - x)\, L_n + n L_{n-1} = 0 \tag{9.79}$$

$$xL_n' = nL_n - nL_{n-1} \tag{9.80}$$

$$\phantom{xL_n'} = L_{n+1} - (n + 1 - x)\, L_n \tag{9.81}$$

### Polinomios de Laguerre asociados

En muchas aplicaciones, principalmente en la teoría cuántica, se requiere de los **polinomios de Laguerre asociados** definidos por

$$L_n^m(x) = (-1)^m \frac{d^m}{dx^m}[L_{n+m}(x)] \tag{9.82}$$

Y de las ecuaciones (9.73) y (9.82) se tiene

$$L_n^m(x) = \sum_{k=0}^{n} \frac{(-1)^k (n+m)!}{(n-k)!(m+k)!k!} x^k \tag{9.83}$$

Derivando la relación (9.78) $m$ veces con respecto a $x$ se obtiene la función generadora de los polinomios de Laguerre asociados

$$\frac{1}{(1-t)^{m+1}} \exp\left(\frac{-xt}{1-t}\right) = \sum_{n=0}^{\infty} L_n^m(x)\, t^n, \quad |t| < 1 \tag{9.84}$$

Las relaciones de recurrencia se establecen fácilmente a partir de la función generadora o por medio de la diferenciación de las relaciones de recurrencia de los polinomios de Laguerre. Dentro de la gran variedad de éstas, se tiene

$$(n+1) L_{n+1}^m(x) = (2n+m+1-x) L_n^m(x) - (n+m) L_{n-1}^m(x) \tag{9.85}$$

$$x[L_n^m(x)]' = nL_n^m(x) - (n+m) L_{n-1}^m(x) \tag{9.86}$$

A partir de éstas, o derivando $m$ veces la ecuación diferencial de Laguerre, se obtiene la ecuación de Laguerre asociada

$$\begin{aligned} xu'' + (m+1-x)u' + nu &= 0 \\ x[L_n^m(x)]'' + (m+1-x)[L_n^m(x)]' + nL_n^m(x) &= 0 \end{aligned} \tag{9.87}$$

La ecuación (9.87) no es autoadjunta, pero se puede pasar a la forma autoadjunta mediante la multiplicación por $x^m e^{-x}$, quedando

$$\left(x^{m+1} e^{-x} u'\right)' + \lambda x^m e^{-x} u = 0 \tag{9.88}$$

$$p(x) = x^{m+1} e^{-x}, \quad \rho(x) = x^m e^{-x}, \quad q(x) = 0, \quad 0 = a \leq x < b = \infty \tag{9.89}$$

Para el eigenvalor $\lambda_n = n$, una solución particular de (9.88) es el polinomio de Laguerre asociado $L_n^m(x)$. Las soluciones $L_n^m(x)$ y $L_l^m(x)$, correspondientes a diferentes eigenvalores $\lambda_n = n$ y $\lambda_l = l$, son ortogonales con el peso $\rho(x) = x^m e^{-x}$

$$\begin{aligned} \left(L_n^m(x), L_l^m\right)_\rho &= \int_0^\infty L_n^m(x) L_l^m(x) x^m e^{-x} dx \\ &= \frac{(n+m)!}{n!} \delta_{nl}, \quad n, l = 0, 1, 2, \ldots \end{aligned} \tag{9.90}$$

La representación de Rodrigues para el polinomio de Laguerre asociado es

$$n! L_n^m(x) = e^x x^{-m} \frac{d^n}{dx^n}\left(e^{-x} x^{n+m}\right) \tag{9.91}$$

### 9.5.4. Función gamma $\Gamma(z)$

Para fines prácticos, basta saber que $\Gamma(z)$ se define mediante la integral

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt, \quad \operatorname{Re} z > 0 \tag{9.92}$$

en donde $z$ es una variable compleja, en general. Esta función es regular para toda $z$, excepto $z = 0, -1, -2, \ldots$, donde tiene polos simples con residuos $\operatorname{Res}\Gamma(z = -m) = (-1)^m/m!$, $m = 0, 1, 2, \ldots$

Al integrar por partes se obtiene

$$\Gamma(z+1) = \int_0^\infty e^{-t} t^z dt = -e^{-t} t^z\Big|_0^\infty + z\int_0^\infty e^{-t} t^{z-1} dt$$

La primera expresión de la derecha es cero y la integral que aparece es nada más que $\Gamma(z)$; esto conduce a la relación básica

$$\Gamma(z+1) = z\Gamma(z) \tag{9.93}$$

Como

$$\Gamma(1) = \int_0^\infty e^{-t} dt = 1$$

con base en (9.93) se concluye que

$$\Gamma(n+1) = n! \qquad (n = 0, 1, 2, \ldots) \tag{9.94}$$

A continuación presentamos sin demostración algunas fórmulas útiles (las demostraciones se pueden encontrar, por ejemplo, en [15]). Puede comprobarse que es válida la igualdad

$$\Gamma(z)\,\Gamma(1-z) = \frac{\pi}{\operatorname{sen} \pi z} \tag{9.95}$$

de donde, en particular, se tiene

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \tag{9.96}$$

Una aproximación de la función gamma para grandes valores de $|z|$ está dada por la **fórmula asintótica de Stirling**,

$$\ln\Gamma(z) \sim \left(z - \frac{1}{2}\right)\ln z - z + \frac{1}{2}\ln 2\pi, \quad \text{o}$$

$$\Gamma(z) \sim \sqrt{2\pi} z^{(z-1/2)} e^{-z} \tag{9.97}$$

Para $z = n + 1$, de la fórmula anterior se tiene

$$n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n} \tag{9.98}$$

La función $B(z,\zeta)$ de dos variables $z$ y $\zeta$, definida por la integral

$$B(z,\zeta) = \int_0^1 x^{z-1}(1-x)^{\zeta-1}\,dx \tag{9.99}$$

$$= \int_0^\infty \frac{t^{z-1}}{(t+1)^{z+\zeta}}dt \tag{9.100}$$

donde $\mathrm{Re}\, z > 0$, $\mathrm{Re}\,\zeta > 0$ y $x = t/(1-t)$, recibe el nombre de **función beta de Euler**. Es fácil probar que $B(z,\zeta) = B(\zeta,z)$ haciendo el cambio de variable $x = 1-y$. Puede comprobarse que también es posible representar la función beta en términos de la función gamma

$$B(z,\zeta) = \frac{\Gamma(z)\,\Gamma(\zeta)}{\Gamma(z+\zeta)} \tag{9.101}$$

### 9.5.5. Funciones de Bessel

Las funciones de Bessel aparecen en una amplia variedad de problemas físicos, principalmente relacionados con el método de separación de variables en la ecuación de onda, de Helmholtz o de Schrodinger, en coordenadas cilíndricas o esféricas. Dichos problemas se reducen a un problema de eigenvalores $\lambda$ descrito por la siguiente ecuación

$$(\tilde{x}u')' + \left(\lambda\tilde{x} - \frac{\nu^2}{\tilde{x}}\right)u = 0 \tag{9.102}$$

$$p(\tilde{x}) = \tilde{x}, \quad \rho(\tilde{x}) = \tilde{x}, \quad q(\tilde{x}) = \frac{\nu^2}{\tilde{x}}; \quad 0 = a < \tilde{x} < b = R \tag{9.103}$$

donde $\nu^2$ es un parámetro dado. Se sabe de la teoría de Sturm-Liouville que las condiciones en la frontera determinan los eigenvalores y proporcionan la ortogonalidad de eigenfunciones. Sin embargo, para los fines actuales no vamos a especificar estas condiciones, dejando la cuestión de la ortogonalidad para problemas más específicos.

Haciendo el cambio de variable $\sqrt{\lambda}\tilde{x} = x$ y denotando $y(x) = u\left(x/\sqrt{\lambda}\right)$, la ecuación (9.102) puede escribirse en la forma

$$\frac{d^2y}{dx^2} + \frac{1}{x}\frac{dy}{dx} + \left(1 - \frac{\nu^2}{x^2}\right)y = 0 \tag{9.104}$$

$$x^2\frac{d^2y}{dx^2} + x\frac{dy}{dx} + \left(x^2 - \nu^2\right)y = 0 \tag{9.105}$$

conocida como la ecuación de Bessel. La solución de la ecuación (9.104) se busca por el método de Frobenius en la forma de serie

$$y(x) = x^r \sum_{k=0}^\infty c_k x^k, \quad c_0 \neq 0 \tag{9.106}$$

Al sustituir esta expresión y sus derivadas en la ecuación de Bessel, se tiene

$$\begin{aligned} 0 &= \sum_{k=0}^{\infty} (k+r)(k+r-1)\, c_k x^{k+r} + \sum_{k=0}^{\infty} (k+r)\, c_k x^{k+r} + \\ & \sum_{k=0}^{\infty} c_k x^{k+r+2} - \nu^2 \sum_{k=0}^{\infty} c_k x^{k+r} \end{aligned}$$

Al igualar a cero la suma de los coeficientes de las potencias $x^{s+r}$, se obtienen las relaciones de recurrencia

$$\left[r(r-1)+r-\nu^2\right] c_0 = 0, \quad s=0 \tag{9.107}$$

$$\left[(r+1)\,r+(r+1)-\nu^2\right] c_1 = 0, \quad s=1 \tag{9.108}$$

$$\left[(s+r)(s+r-1)+(s+r)-\nu^2\right] c_s + c_{s-2} = 0, \quad s=2,3,\ldots \tag{9.109}$$

De la ecuación (9.107) se obtiene la ecuación indicial

$$(r+\nu)(r+\nu) = 0 \tag{9.110}$$

que tiene las raíces $r_1 = \nu$ y $r_2 = -\nu$. Se sabe que en el método de Frobenius se distinguen tres casos (véase la sección 9.4):

1. Raíces distintas $r_1$ y $r_2$ que no difieren en un entero, es decir, $\nu$ no es entero;

2. Raíz doble, que se refiere a $\nu = 0$;

3. Raíces distintas $r_1$ y $r_2$ que difieren en un entero; se obtienen cuando $\nu$ es un entero o semientero.

Esto conduce a que las soluciones generales de la ecuación de Bessel, para cada uno de estos casos, sean diferentes.

**Funciones de Bessel de primera clase**

En todo caso, una solución particular es de la forma (9.106). Para el valor de $r = \nu$, las ecuaciones (9.108) y (9.109) toman la forma

$$(2\nu+1)\, c_1 = 0 \tag{9.111}$$

$$s\,(s+2\nu)\, c_s + c_{s-2} = 0, \quad s=2,3,\ldots \tag{9.112}$$

Suponemos que $\nu$ no es un número semientero negativo, $\nu \neq -1/2, -3/2, \ldots, -(2m+1)/2, \ldots$; en el caso opuesto, el razonamiento que sigue es válido para $r_2 = -\nu$. De la ecuación (9.111) se tiene que $c_1 = 0$. Dado que $c_1 = 0$, de la ecuación (9.112) se deduce que todos los coeficientes de índice impar son ceros, $c_{2m+1} = 0$ $(m = 1, 2, 3, \ldots)$. Para los coeficientes de índice par se obtiene la fórmula de recurrencia

$$c_{2m} = -\frac{c_{2m-2}}{2^2 m\,(\nu+m)}, \quad m = 1, 2, 3, \ldots \tag{9.113}$$

y, a partir de ésta se pueden determinar sucesivamente todos los coeficientes de índice par. El coeficiente $c_0$ permanece arbitrario; por conveniencia, se acostumbra poner

$$c_0 = \frac{1}{2^\nu \Gamma(\nu+1)} \tag{9.114}$$

De las ecuaciones (9.113) y(9.114) se tiene

$$c_{2m} = \frac{(-1)^m}{2^{2m+\nu} m! \Gamma(\nu+m+1)}, \quad m = 1, 2, 3, \ldots \tag{9.115}$$

Al sustituir estos coeficientes en (9.106) y recordando que $c_{2m+1} = 0$ ($m = 1, 2, 3, \ldots$), se obtiene una solución particular de la ecuación de Bessel, la cual se denota por $J_\nu(x)$:

$$J_\nu(x) = x^\nu \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m}}{2^{2m+\nu} m! \Gamma(\nu+m+1)} \tag{9.116}$$

Esta solución de (9.104), conocida como **función de Bessel de primera clase y de orden** $\nu$, tiene la forma de (9.106). La serie converge para toda $x$, lo cual puede verificarse mediante la prueba de la razón.

La solución particular $J_\nu(x)$ de la ecuación de Bessel existe para cualquier $\nu^2$. Reemplazando $\nu$ por $-\nu$ en (9.116), se tiene

$$J_{-\nu}(x) = x^{-\nu} \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m}}{2^{2m-\nu} m! \Gamma(m-\nu+1)} \tag{9.117}$$

Puesto que la ecuación de Bessel contiene a $\nu^2$, tanto $J_\nu(x)$ como $J_{-\nu}(x)$ son soluciones de ésta. Si $\nu$ no es un entero ni cero, estas soluciones son linealmente independientes, debido a que las potencias de $x$ en (9.116) no se repiten en (9.117). Entonces, si $\nu$ no es un entero ni cero, la solución general de la ecuación de Bessel, para toda $x \neq 0$, es

$$y(x) = C_1 J_\nu(x) + C_2 J_{-\nu}(x) \tag{9.118}$$

donde $C_1$ y $C_2$ son dos constantes arbitrarias.

Si $\nu$ es un entero o cero, la función (9.118) no es una solución general. Para $\nu = n$ entero, las funciones $J_\nu(x)$ y $J_{-\nu}(x)$ son linealmente dependientes, porque

$$J_{-n}(x) = (-1)^n J_n(x) = J_n(-x), \quad n = 0, 1, 2, \ldots \tag{9.119}$$

Para comprobar la última relación, se sustituye $\nu = n \geq 0$ en (9.117). Dado que la función gamma $\Gamma(z)$ tiene polos simples en $z = 0, -1, -2, \ldots$, los primeros $n$ coeficientes de la serie (9.117) son ceros y la suma se inicia con $m = n$. Ya que

$\Gamma(m-n+1) = (m-n)!$, se obtiene

$$\begin{aligned} J_{-n}(x) &= x^{-n} \sum_{m=n}^{\infty} \frac{(-1)^m x^{2m}}{2^{2m-n} m!\,(m-n)!} \\ &= \sum_{s=0}^{\infty} \frac{(-1)^{n+s} x^{2s+n}}{2^{2s+n}(s+n)!s!} = (-1)^n J_n(x) \\ &= (-x)^n \sum_{s=0}^{\infty} \frac{(-1)^s (-x)^{2s}}{2^{2s+n}(s+n)!s!} = J_n(-x) \end{aligned}$$

en donde $m-n=s$ y $m=n+s$. Esto completa la demostración de (9.119).

La función

$$G(x,h) = \exp\left(\frac{x}{2}\left(h-\frac{1}{h}\right)\right) \tag{9.120}$$

se conoce como **función generadora** de las funciones de Bessel de primera clase, en virtud de que su desarrollo en serie de potencias de $h$ es

$$\begin{aligned} \exp\left(\frac{x}{2}\left(h-\frac{1}{h}\right)\right) &= \sum_{n=-\infty}^{\infty} J_n(x)\, h^n \qquad (9.121) \\ &= J_0(x) + \sum_{n=1}^{\infty} J_n(x)\left(h^n + (-h)^{-n}\right) \end{aligned}$$

La función generadora es muy útil para establecer varias relaciones de recurrencia y la representación integral de las funciones de Bessel.

Una forma particularmente útil y poderosa de tratar las funciones de Bessel utiliza la representación integral

$$J_n(x) = \frac{1}{\pi}\int_0^{\pi} \cos(n\theta - x\,\mathrm{sen}\,\theta)\,d\theta, \quad n = 0,1,2,\ldots \tag{9.122}$$

Para comprobar la relación anterior, sustituimos $h = e^{i\theta}$ en (9.121)

$$\begin{aligned} \exp(ix\,\mathrm{sen}\,\theta) &= J_0(x) + \sum_{n=1}^{\infty} J_n(x)\left(e^{in\theta} + (-1)^n e^{-in\theta}\right) \\ &= J_0(x) + 2\sum_{m=1}^{\infty} J_{2m}(x)\cos 2m\theta + \\ &\quad 2i\sum_{m=1}^{\infty} J_{2m-1}(x)\,\mathrm{sen}\,(2m-1)\,\theta \end{aligned}$$

en donde hemos utilizado las relaciones trigonométricas conocidas, $\cos\theta = \left(e^{i\theta} + e^{-i\theta}\right)/2$ y $\mathrm{sen}\left(e^{i\theta} - e^{-i\theta}\right)/2i$. Igualando las partes real e imaginaria,

respectivamente, se tiene

$$\cos(x \operatorname{sen}\theta) = J_0(x) + 2\sum_{m=1}^{\infty} J_{2m}(x)\cos 2m\theta \tag{9.123}$$

$$\operatorname{sen}(x \operatorname{sen}\theta) = 2\sum_{m=1}^{\infty} J_{2m-1}(x)\operatorname{sen}(2m-1)\theta \tag{9.124}$$

Utilizando la ortogonalidad de las funciones seno y coseno

$$\int_0^{\pi} \cos n\theta \cos m\theta d\theta = \frac{\pi}{2}\delta_{nm}$$
$$\int_0^{\pi} \operatorname{sen} n\theta \operatorname{sen} m\theta d\theta = \frac{\pi}{2}\delta_{nm}$$

en donde $n, m = 1, 2, \ldots$, de las ecuaciones (9.123) y (9.124) se obtiene

$$\frac{1}{\pi}\int_0^{\pi} \cos(x \operatorname{sen}\theta)\cos n\theta d\theta = \begin{cases} J_n(x), & \text{si } n \text{ es par} \\ 0, & \text{si } n \text{ es impar} \end{cases}$$
$$\frac{1}{\pi}\int_0^{\pi} \operatorname{sen}(x \operatorname{sen}\theta)\operatorname{sen} n\theta d\theta = \begin{cases} 0, & \text{si } n \text{ es par} \\ J_n(x), & \text{si } n \text{ es impar} \end{cases}$$

Sumando estas dos ecuaciones por partes se tiene

$$\begin{aligned} J_n(x) &= \frac{1}{\pi}\int_0^{\pi} [\cos(x \operatorname{sen}\theta)\cos n\theta + \operatorname{sen}(x \operatorname{sen}\theta)\operatorname{sen} n\theta]\, d\theta \\ &= \frac{1}{\pi}\int_0^{\pi} \cos(n\theta - x\operatorname{sen}\theta)\, d\theta, \quad n = 0, 1, 2, \ldots \end{aligned}$$

que completa la comprobación de (9.33).

**Problema.** *Utilice la ecuación (9.116) y la relación* $\Gamma(\nu+1) = \nu\Gamma(\nu)$ *para demostrar la relación de recurrencia*

$$[x^{\nu} J_{\nu}(x)]' = x^{\nu} J_{\nu-1}(x) \tag{9.125}$$

*De manera similar, demuestre que*

$$[x^{-\nu} J_{\nu}(x)]' = -x^{-\nu} J_{\nu+1}(x) \tag{9.126}$$

**Problema.** *A partir de las ecuaciones (9.125) y (9.126) obtenga las relaciones de recurrencia*

$$J_{\nu-1}(x) + J_{\nu+1}(x) = \frac{2\nu}{x} J_{\nu}(x) \tag{9.127}$$

$$J_{\nu-1}(x) - J_{\nu+1}(x) = 2J'_{\nu}(x) \tag{9.128}$$

**Problema.** *Utilizando la ecuación (9.116) y la relación* $\Gamma(1/2) = \sqrt{\pi}$, *demuestre que*

$$J_{1/2}(x) = \sqrt{\frac{2}{\pi x}}\operatorname{sen} x, \quad J_{-1/2}(x) = \sqrt{\frac{2}{\pi x}}\cos x \tag{9.129}$$

*A partir de este resultado y (9.127), demuestre que las funciones de Bessel* $J_\nu(x)$ *de orden semientero* $\nu = \pm 1/2, \pm 3/2, \pm 5/2, \ldots$ *son funciones elementales.*

**Problema.** *Utilizando (9.127) y (9.129), demuestre que*

$$\begin{aligned} J_{3/2}(x) &= \sqrt{\frac{2}{\pi x}}\left(\frac{\operatorname{sen} x}{x} - \cos x\right) \\ J_{-3/2}(x) &= \sqrt{\frac{2}{\pi x}}\left(\frac{\cos x}{x} + \operatorname{sen} x\right) \end{aligned}$$

Cabe mencionar que las relaciones de recurrencia, ecuaciones (9.125), (9.126), (9.127) y (9.128), para las funciones de Bessel $J_n(x)$ de orden entero, se obtienen también mediante la derivación de la función generadora, ecuación (9.121).

**Problema.** *Para los valores pequeños de* $x$, *verifique que el término principal del desarrollo de* $J_\nu(x)$ *es*

$$J_\nu(x) \approx \frac{x^\nu}{2^\nu \Gamma(\nu+1)} - \cdots \tag{9.130}$$

*quedando*

$$\begin{aligned} \lim_{x\to+0} J_0(x) &= 1 \\ \lim_{x\to+0} J_\nu(x) &= 0, \quad \nu > 0 \end{aligned}$$

## Ortogonalidad de las funciones de Bessel

Como se mostró al inicio de esta sección, la ecuación de Bessel (9.105) está estrechamente relacionada con la ecuación autoadjunta de Sturm-Liouville (9.102). Por lo tanto, para cada $\nu$ fijo, según el teorema de Sturm-Liouville, las soluciones de la ecuación de Bessel $J_\nu\left(\sqrt{\lambda}\tilde{x}\right)$ correspondientes a diferentes eigenvalores $\lambda$, son ortogonales sobre el intervalo $0 \le \tilde{x} \le R$ con respecto a la función peso $\rho(\tilde{x}) = \tilde{x}$, donde $x = \sqrt{\lambda}\tilde{x}$. Ya que $p(\tilde{x}) = \tilde{x}$ (véase la ecuación (9.103)) es cero en $\tilde{x} = 0$, la condición en la frontera $\tilde{x} = 0$ no tiene importancia para la ortogonalidad de las eigenfunciones (véase la sección 9.1). A su vez, la condición en la frontera $\tilde{x} = R$ nos proporciona la ecuación del tipo (9.2)

$$\alpha J_\nu\left(\sqrt{\lambda}R\right) + \beta J_\nu'\left(\sqrt{\lambda}R\right) = 0 \tag{9.131}$$

en donde $J'_\nu\left(\sqrt{\lambda}R\right) = dJ_\nu(x)/dx$ para $x = \sqrt{\lambda}R$. Resolviendo la última con respecto a $\lambda$ para $\nu$ fijo, se tiene un conjunto de eigenvalores del problema de Sturm-Liouville

$$\lambda_{\nu k}, \quad k = 1, 2, 3, \ldots \tag{9.132}$$

Las eigenfunciones correspondientes a diferentes eigenvalores son ortogonales

$$\int_0^R J_\nu\left(\sqrt{\lambda_{\nu k}}\tilde{x}\right) J_\nu\left(\sqrt{\lambda_{\nu m}}\tilde{x}\right)\tilde{x}d\tilde{x} = \frac{R^2}{2}\left[J_{\nu+1}\left(\sqrt{\lambda_{\nu k}}R\right)\right]^2 \delta_{km} \tag{9.133}$$

para $k, m = 1, 2, \ldots$ Nótese que se tiene una infinidad de conjuntos ortogonales de funciones de Bessel y que cada conjunto corresponde a uno de los valores fijos de $\nu$.

El conjunto de funciones de Bessel $\left\{J_\nu\left(\sqrt{\lambda_{\nu k}}\tilde{x}\right) : \ k = 1, 2, \ldots; \nu \text{ fijo}\right\}$ es completo, entonces cualquier función $f(\tilde{x})$ de comportamiento adecuado se puede desarrollar en una serie de Fourier-Bessel

$$f(\tilde{x}) = \sum_{k=1}^{\infty} c_{\nu k} J_\nu\left(\sqrt{\lambda}\tilde{x}\right), \quad 0 \leq \tilde{x} \leq R, \quad \nu > -1 \tag{9.134}$$

con coeficientes

$$c_{\nu k} = \frac{2}{R^2\left[J_{\nu+1}\left(\sqrt{\lambda_{\nu k}}R\right)\right]^2}\int_0^R f(\tilde{x}) J_\nu\left(\sqrt{\lambda_{\nu m}}\tilde{x}\right)\tilde{x}d\tilde{x} \tag{9.135}$$

### Funciones de Bessel de segunda clase (funciones de Neumann)

Para $\nu = n$ entero, las funciones $J_n(x)$ y $J_{-n}(x)$ son linealmente dependientes (véase la ecuación (9.119)); por tanto, no forman una base de soluciones. Con el fin de obtener una segunda solución linealmente independiente, cuando $\nu$ es un entero, seguimos con el método de Frobenius.

Se empieza por $\nu = 0$. Para construir una solución parcial de la ecuación de Bessel, que es linealmente independiente de $J_0(x)$, se puede usar el método de Frobenius para el caso de una raíz doble de la ecuación indicial (sección 9.3). En este caso, la ecuación de Bessel se puede escribir

$$xy'' + y' + xy = 0 \tag{9.136}$$

y la solución deseada debe tener la forma (véanse las ecuaciones (9.26) y (9.30))

$$y_{2,0}(x) = J_0(x)\ln x + \sum_{k=1}^{\infty} A_k x^k \tag{9.137}$$

Sustituyendo $y_2(x)$ y sus derivadas en (9.136) y usando el hecho de que $J_0(x)$ es la solución de (9.136), se obtiene la ecuación

$$2J'_0 + \sum_{k=1}^{\infty} k(k-1)A_k x^{k-1} + \sum_{k=1}^{\infty} kA_k x^{k-1} + \sum_{k=1}^{\infty} A_k x^{k+1} = 0$$

A partir de (9.116) se representa $J_0'$ en la forma de serie de potencias

$$J_0'(x) = \sum_{k=1}^{\infty} \frac{(-1)^k x^{2k-1}}{2^{2k-1} k! (k-1)!}$$

Al introducir esta serie de potencias en la ecuación anterior, se tiene

$$\sum_{k=1}^{\infty} \frac{(-1)^k x^{2k-1}}{2^{2k-2} k! (k-1)!} + \sum_{k=1}^{\infty} k^2 A_k x^{k-1} + \sum_{k=1}^{\infty} A_k x^{k+1} = 0 \tag{9.138}$$

La primera serie de esta ecuación contiene únicamente potencias impares. Por tanto, las potencias pares e impares se consideran por separado.

El coeficiente de la potencia $x^0 = 1$ es $A_1$ y, en consecuencia, $A_1 = 0$. Igualando a cero los coeficientes de las potencias $x^{2s}$, se tiene

$$(2s+1)^2 A_{2s+1} + A_{2s-1} = 0, \quad s = 1, 2, \ldots \tag{9.139}$$

Tomando en cuenta que $A_1 = 0$, de la relación anterior se concluye que $A_{2s-1} = 0$ $(s = 1, 2, \ldots)$.

Igualando a cero los coeficientes de las potencias impares $x^{2s-1}$, se tiene

$$\frac{(-1)^s}{2^{2s-2} s! (s-1)!} + (2s)^2 A_{2s} + A_{2s-2} = 0, \quad s = 1, 2, \ldots \tag{9.140}$$

Para $s = 1$ queda como

$$-1 + 4A_2 + A_0 = 0 \tag{9.141}$$

quedando $A_0$ como una constante arbitraria. Por simplicidad, elegimos $A_0 = 0$. Entonces, de las ecuaciones (9.140) y (9.141), se tiene

$$A_{2s} = \frac{(-1)^{s-1} h_s}{2^{2s} (s!)^2}, \quad s = 1, 2, \ldots \tag{9.142}$$

en donde se aplica la notación

$$h_s = 1 + \frac{1}{2} + \cdots + \frac{1}{s} \tag{9.143}$$

Sustituyendo los coeficientes $A_k$ en la ecuación (9.137) se obtiene

$$y_{2,0}(x) = J_0(x) \ln x + \sum_{k=1}^{\infty} \frac{(-1)^{k-1} h_k}{2^{2k} (k!)^2} x^{2k} \tag{9.144}$$

Dado que $J_0(x)$ y $y_2(x)$ son funciones linealmente independientes, forman una base de soluciones de la ecuación (9.136). Sin embargo, se acostumbra elegir como la segunda solución una combinación lineal de $J_0(x)$ y $y_2(x)$ conocida

como **función de Bessel de segunda clase** de orden cero o **función de Neumann** de orden cero

$$\begin{aligned} Y_0(x) &= \frac{2}{\pi}\left[y_{2,0}(x) + (\gamma - \ln 2) J_0(x)\right] \qquad (9.145) \\ &= \frac{2}{\pi}\left[J_0(x)\left(\ln\frac{x}{2} + \gamma\right) + \sum_{k=1}^{\infty} \frac{(-1)^{k-1} h_k}{2^{2k}(k!)^2} x^{2k}\right] \end{aligned}$$

en donde $\gamma$ es la llamada **constante de Euler**, la cual se define como el límite

$$\gamma = \lim_{s\to\infty}\left(\sum_{k=1}^{s}\frac{1}{k} - \ln s\right) = 0.57721566490 \qquad (9.146)$$

Dado que $J_0(x) \to 1$ cuando $x \to +0$, para $x > 0$ pequeña, es fácil ver que la función $Y_0(x)$ se comporta como $Y_0(x) \approx 2\pi^{-1}(\ln(x/2) + \gamma)$ y $Y_0(x) \to -\infty$ cuando $x \to +0$.

Una solución parcial de la ecuación de Bessel con $\nu = n = 1, 2, \ldots$ entero, que es linealmente independiente de $J_n(x)$, se puede obtener por el método de Frobenius para el caso de raíces distintas de la ecuación indicial que difieren en un entero. Resulta que, en estos casos, la solución contiene también un término logarítmico, quedando de la forma (véase la ecuación (9.32))

$$y_{2,-n}(x) = K_{2n} J_n(x) \ln x + x^{-n} \sum_{k=0}^{\infty} A_k x^k, \quad x > 0 \qquad (9.147)$$

La situación todavía no es completamente satisfactoria, debido a que la segunda solución se define de modo diferente, dependiendo de si el orden $\nu$ es entero o no. Con el fin de proporcionar la uniformidad del formalismo, resulta deseable adoptar una forma de la segunda solución que sea válida para toda $\nu$. Por esta razón se introduce una segunda solución estándar $Y_n(x)$, definida para toda $\nu$ por la fórmula

$$Y_\nu(x) = \frac{J_\nu(x)\cos\nu\pi - J_{-\nu}(x)}{\operatorname{sen}\nu\pi} \qquad (9.148)$$

$$Y_n(x) = \lim_{\nu\to n} Y_\nu(x), \quad n = 0, 1, 2, \ldots \qquad (9.149)$$

La función $Y_\nu(x)$ se conoce como **función de Bessel de segunda clase** de orden $\nu$ o **función de Neumann** de orden $\nu$.

Para $\nu$ no entero, $Y_\nu(x)$ evidentemente satisface la ecuación de Bessel, ya que es una combinación lineal de las soluciones linealmente independientes conocidas, $J_\nu(x)$ y $J_{-\nu}(x)$. Para $\nu = n$ entero, aplicando la fórmula (9.119), se ve que la expresión (9.148) se vuelve indeterminada. Sin embargo, evaluando $Y_\nu(x)$ mediante la regla de L'Hopital para las formas indeterminadas, se obtiene

$$\begin{aligned} Y_n(x) &= \left.\frac{\partial\left(J_\nu(x)\cos\nu\pi - J_{-\nu}(x)\right)/\partial\nu}{\partial(\operatorname{sen}\nu\pi)/\partial\nu}\right|_{\nu=n} \\ &= \frac{1}{\pi}\left[\frac{\partial J_\nu(x)}{\partial\nu} - (-1)^\nu \frac{\partial J_{-\nu}(x)}{\partial\nu}\right]_{\nu=n} \qquad (9.150) \end{aligned}$$

Un desarrollo en serie proporciona el resultado

$$\begin{aligned} Y_n(x) &= \frac{2}{\pi} J_n(x) \left( \ln \frac{x}{2} + \gamma \right) + \frac{x^n}{\pi} \sum_{k=0}^{\infty} \frac{(-1)^{k-1} (h_k + h_{k+n})}{2^{2k+n} k! (k+n)!} x^{2k} \\ &- \frac{x^{-n}}{\pi} \sum_{k=0}^{n-1} \frac{(n-k-1)!}{2^{2k-n} k!} x^{2k} \end{aligned} \tag{9.151}$$

donde se utilizó la fórmula $(\partial/\partial\nu)\, x^{\nu} = x^{\nu} \ln x$. Para $\nu = 0$, la última suma que aparece en (9.151) se debe reemplazar por 0 y (9.151) toma la forma (9.145). Además, se puede demostrar que

$$Y_{-n}(x) = (-1)^n Y_n(x) \tag{9.152}$$

Para verificar que $Y_n(x)$ satisface la ecuación de Bessel para $\nu = n$ entero, se puede proceder como sigue. Derivando la ecuación de Bessel para $J_{\pm\nu}(x)$ con respecto a $\nu$, se tiene

$$x^2 \frac{d^2}{dx^2} \left( \frac{\partial J_{\pm\nu}}{\partial \nu} \right) + x \frac{d}{dx} \left( \frac{\partial J_{\pm\nu}}{\partial \nu} \right) + \left( x^2 - \nu^2 \right) \frac{\partial J_{\pm\nu}}{\partial \nu} = 2\nu J_{\pm\nu}$$

Multiplicando la ecuación de $J_{-\nu}$ por $(-1)^{\nu}$, restando de la ecuación para $J_{\nu}$ (como se sugiere mediante la ecuación (9.150)) y considerando el límite $\nu \to n$, se obtiene

$$x^2 \frac{d^2}{dx^2} Y_n(x) + x \frac{d}{dx} Y_n(x) + \left( x^2 - \nu^2 \right) Y_n(x) = \frac{2n}{\pi} \left[ J_n - (-1)^n J_{-n} \right] = 0$$

porque $J_{-n} = (-1)^n J_n$. Por tanto, $Y_n(x)$ corresponde a una solución de la ecuación de Bessel. Además, el desarrollo en serie de $Y_n(x)$, ecuación (9.151), contiene un término logarítmico; por consiguiente, $J_n(x)$ y $Y_n(x)$ son soluciones linealmente independientes. Entonces, la solución general de la ecuación de Bessel para cualquier $\nu$ es

$$y(x) = C_1 J_n(x) + C_2 Y_n(x) \tag{9.153}$$

**Problema.** *Demuestre que las funciones de Neumann $Y_n(x)$, con $\nu = n$ un número entero, satisfacen las relaciones de recurrencia*

$$Y_{n-1}(x) + Y_{n+1}(x) = \frac{2n}{x} Y_n(x) \tag{9.154}$$

$$Y_{n-1}(x) - Y_{n+1}(x) = 2Y_n'(x) \tag{9.155}$$

*Estas relaciones se pueden demostrar mediante la diferenciación de las relaciones de recurrencia para $J_{\nu}$ o mediante el uso de la forma de límite de $Y_{\nu}(x)$.*

**Problema.** *Demuestre que $Y_{-n}(x) = (-1)^n Y_n(x)$.*

**Problema.** *Demuestre que $Y_0'(x) = -Y_1(x)$.*

**Problema.** *Para valores pequeños de $x$, verifique los términos principales de los desarrollos*

$$Y_0(x) \approx \frac{2}{\pi}(\ln x + \gamma - \ln 2) \tag{9.156}$$

$$Y_\nu(x) \approx -\frac{(\nu-1)!}{\pi}\left(\frac{2}{x}\right)^\nu, \quad \nu > 0 \tag{9.157}$$

*quedando*

$$\lim_{x\to+0} Y_\nu(x) = -\infty, \quad \nu \geq 0 \tag{9.158}$$

De la ecuación anterior se observa que $Y_\nu(x)$ diverge cuando $x \to +0$. Esto implica que cualquier condición limitante con un requisito de que la función sea finita en el origen excluye de manera automática a $Y_\nu(x)$ de la solución general (9.153). En ausencia de tal requerimiento, se debe considerar a $Y_\nu(x)$.

## Funciones de Bessel de tercera clase (funciones de Hankel)

Por una necesidad práctica, para tener soluciones de la ecuación de Bessel que sean complejas para los valores reales de $x$, con frecuencia se utilizan las soluciones

$$H_\nu^{(1)}(x) = J_\nu(x) + iY_\nu(x) \tag{9.159}$$

$$H_\nu^{(2)}(x) = J_\nu(x) - iY_\nu(x) \tag{9.160}$$

Estas funciones linealmente independientes se conocen como **funciones de Bessel de tercera clase** de orden $\nu$ o primera y segunda **funciones de Hankel** de orden $\nu$. Ya que las funciones de Hankel son combinaciones lineales de $J_\nu(x)$ y $Y_\nu(x)$, satisfacen las mismas relaciones de recurrencia, ecuaciones (9.127) y (9.128) o ecuaciones (9.154) y (9.155).

Por último, con el fin de resaltar la importancia práctica de las funciones de Bessel se presenta una variedad de ecuaciones diferenciales reducibles a la ecuación de Bessel.

Cuando la ecuación de Helmholtz en coordenadas polares esféricas se separa, la ecuación radial tiene la forma

$$r^2\frac{d^2R}{dr^2} + 2r\frac{dR}{dr} + \left[k^2r^2 - n(n+1)\right]R = 0 \tag{9.161}$$

en donde $k$ proviene de la ecuación de Helmholtz y el número entero $n$ es una constante de separación que a menudo tiene el significado físico de momento angular. El cambio de variable independiente $x = kr$ transforma la ecuación (9.161) en

$$x^2R'' + 2xR' + \left[x^2 - n(n+1)\right]R = 0$$

Mediante la sustitución $R = x^{-1/2}y$, la última ecuación se reduce a la ecuación de Bessel para la función $y$

$$x^2y'' + xy' + \left[x^2 - (n+1/2)^2\right]y = 0 \tag{9.162}$$

cuyas soluciones linealmente independientes son $J_{n+1/2}(x)$ y $Y_{n+1/2}(x)$. Por tanto, soluciones linealmente independientes de (9.161) son

$$R_1(x) = \frac{J_{n+1/2}(kr)}{(kr)^{1/2}}, \quad R_2(x) = \frac{Y_{n+1/2}(kr)}{(kr)^{1/2}}$$

La combinación de funciones anteriores se presenta con bastante frecuencia. Debido a la importancia de las coordenadas esféricas, se acostumbra llamar a las funciones

$$J_n(x) = \sqrt{\frac{\pi}{2x}} J_{n+1/2}(x), \tag{9.163}$$

$$y_n(x) = \sqrt{\frac{\pi}{2x}} Y_{n+1/2}(x) = (-1)^{n+1} \sqrt{\frac{\pi}{2x}} J_{-n-1/2}(x) \tag{9.164}$$

como **funciones de Bessel esféricas**. Utilizando (9.127) y (9.129) es fácil demostrar que las funciones de Bessel esféricas son elementales, es decir, éstas siempre pueden expresarse como sen $x$ y cos $x$ con coeficientes que son polinomios que involucran potencias negativas de $x$.

**Problema.** *Aplique los cambios de variable indicados y encuentre una solución general de las ecuaciones dadas, en términos de las funciones de Bessel:*

$$x^2y'' + xy' + (\lambda^2x^2 - \nu^2)\, y = 0 \quad (\lambda x = t)$$

$$x^2y'' + 2xy' + (\lambda^2x^2 - n(n+1))\, y = 0 \quad (\lambda x = t)$$

$$x^2y'' + xy' + 4(x^4 - \nu^2)\, y = 0 \quad (x^2 = t)$$

$$4x^2y'' + 4xy' + (x - \nu^2)\, y = 0 \quad \left(x^{1/2} = t\right)$$

$$xy'' + y' + y/4 = 0 \quad \left(x^{1/2} = t\right)$$

$$xy'' - y' + xy = 0 \quad (y = xu)$$

$$x^2y'' + (x^2 + 1/4)\, y = 0 \quad \left(y = x^{1/2}u\right)$$

$$xy'' + (1 + 2n)\, y' + xy = 0 \quad (y = x^{-n}u)$$

$$x^2y'' + (x + 3/4)\, y/4 = 0 \quad \left(y = x^{1/2}u,\ x^{1/2} = t\right)$$

$$x^2y'' - 3xy' + 4(x^2 - 3)\, y = 0 \quad (y = x^2u,\ x^2 = t)$$

$$y'' + xy = 0 \quad \left(y = x^{1/2}u,\ 2x^{3/2}/3 = t\right)$$

$$x^2y'' + (1 - 2\nu)\, xy' + \nu^2(x^{2\nu} + 1 - \nu^2)\, y = 0 \quad (y = x^{\nu}u,\ x^{\nu} = t)$$

**Nota bibliográfica:** para ampliar los conocimientos sobre los temas de este capítulo, puede consultar [1], [2], [4], [6], [11], [17] y [18].

otras soluciones linealmente independientes son $J_{n+1/2}(kr)$ y $Y_{n+1/2}(kr)$; por lo tanto, soluciones linealmente independientes de (9.161) son

$$R_1(x) = \frac{J_{n+1/2}(kr)}{(kr)^{1/2}}, \qquad R_2(x) = \frac{Y_{n+1/2}(kr)}{(kr)^{1/2}}$$

La combinación de funciones anteriores se presenta con bastante frecuencia. Debido a la importancia de las coordenadas esféricas, se acostumbra definir a las funciones

$$j_n(x) = \sqrt{\frac{\pi}{2x}}\, J_{n+1/2}(x), \qquad y_n(x) = \sqrt{\frac{\pi}{2x}}\, Y_{n+1/2}(x)$$

como **funciones de Bessel esféricas**. Utilizando (9.127) y (9.129), es fácil demostrar que las funciones de Bessel esféricas son elementales, es decir, siempre pueden expresarse como sen $x$ y cos $x$ con coeficientes que son polinomios que involucran potencias negativas de $x$.

**Problema.** Aplique los cambios de variable indicados y encuentre la solución general de las ecuaciones dadas, en términos de las *funciones de Bessel*.

$$x^2y'' + xy' + (\lambda^2x^2 - \nu^2)\,y = 0, \quad (\lambda x = t)$$

$$x^2y'' + 2xy' + (\lambda^2x^2 - n(n+1))\,y = 0, \quad (\lambda x = t)$$

$$x^2y'' + xy' + 4(x^4 - \nu^2)\,y = 0, \quad (x^2 = t)$$

$$4x^2y'' + 4xy' + (x - \nu^2)\,y = 0, \quad (x^{1/2} = t)$$

$$x^2y'' + xy' + (e^{2x} - \nu^2)\,y = 0, \quad (e^x = t)$$

$$xy'' - y' + xy = 0, \quad (y = xu)$$

$$x^2y'' + (x^2 + 1/4)\,y = 0, \quad (y = x^{1/2}u)$$

$$xy'' + (1 + 2n)\,y' + xy = 0, \quad (y = x^{-n}u)$$

$$x^2y'' + (x + 3/4)\,y/4 = 0, \quad (y = x^{1/2}u,\ x^{1/2} = t)$$

$$x^2y'' - 3xy' + 4(x^4 - 3)\,y = 0, \quad (y = x^2u,\ x^2 = t)$$

$$y'' + xy = 0, \quad \left(y = x^{1/2}u,\ \tfrac{2}{3}x^{3/2} = t\right)$$

$$x^2y'' + (1 - 2\alpha)\,xy' + (\beta^2\gamma^2x^{2\gamma} + \alpha^2 - \nu^2\gamma^2)\,y = 0, \quad (y = x^\alpha u,\ \beta x^\gamma = t)$$

**Nota bibliográfica:** para ampliar los conocimientos sobre los temas de este capítulo, puede consultar [1], [2], [4], [6], [11], [17] y [18].

# Bibliografía


[1] G. Arfken, *Métodos matemáticos para físicos*, Academic Press, 1985.

[2] F. Ayres Jr., *Teoría y problemas de ecuaciones diferenciales*, Serie Schaum en español, McGraw-Hill, México, 1978.

[3] St. Barnett, *Matrix methods for engineers and scientists*, McGraw-Hill Book Company (UK) Limited, 1979.

[4] M.L. Boas, *Mathematical methods in physical sciences*, John Wiley and Sons, N.Y., 1966.

[5] W.E. Boyce y R.C. DiPrima, *Ecuaciones diferenciales y problemas con valores en la frontera*, 3 ed., Editorial LIMUSA, México, 1984.

[6] E. Butkov, *Mathematical physics*, Addison-Wesley Publishing Company, Massachusetts, 1968.

[7] H. Eves, *Elementary matrix theory*, Dover Publications, N.Y., 1980.

[8] F.R. Gantmacher, *Teoría de matrices*, Nauka, Moscú, 1967 (*en Ruso*).

[9] F.R. Gantmacher, *The theory of matrices*, vols. 1 y 2, Chelsea, Nueva York, 1959.

[10] F.E. Hohn, *Álgebra de matrices,* 3 ed., Editorial Trillas, México, 1981.

[11] E. Kreyszig, *Matemáticas avanzadas para ingeniería,* vol. 1, Editorial LIMUSA, Grupo Noriega Editores, México, 1997.

[12] E. Kreyszig, *Matemáticas avanzadas para ingeniería,* vol. 2, Editorial LIMUSA, Grupo Noriega Editores, México, 1999.

[13] H. Lass, *Vector and tensor analysis*, McGraw-Hill Kogakusha LTD, 1950.

[14] M. Lipshchutz, *Geometría diferencial,* McGraw-Hill, Serie Schaum en español, 1971.

[15] A.I. Markushevich, *Theory of functions of a complex variable*, Chelsea Publishing Company, N.Y., 1985.

[16] B. Noble y James W. Daniel, *Álgebra lineal aplicada*, Prentice Hall Hispanoamericana, México, 1989.

[17] I.S. Sokolnikoff y R.M. Redheffer, *Mathematics of physics and modern engineering*, McGraw-Hill, 1966.

[18] V.S. Vladimirov, *Ecuaciones de física matemática*, Nauka, Moscú, 1967 (*en Ruso*).

# Índice

L[illegible] EDICIÓN, COMPOSICIÓN, DISEÑO E IMPRESIÓN DE ESTA OBRA FUERON REALIZADOS
BAJO LA SUPERVISIÓN DE GRUPO NORIEGA EDITORES
BALDERAS 95, COL. CENTRO. MÉXICO, D.F. C.P. 06040
1205955001002508DP9200IE

En nuestros tiempos de "computación total", existe la creencia bastante común de que los problemas matemáticos aplicados en su mayoría absoluta se resuelven por computadora. Sin embargo, la complejidad de problemas a tratar en las ciencias exactas y naturales requiere la aplicación de conceptos y métodos de las Matemáticas cada vez más complicados y poderosos. Es por tanto, el objetivo principal de este libro de texto, que el estudiante se familiarice profundamente con los conceptos matemáticos fundamentales. El texto pretende abarcar los temas indispensables en matemáticas para los estudios de Física y áreas afines. El propósito de los *Fundamentos de Métodos Matemáticos para Física e Ingeniería* es dar a los estudiantes bases sólidas en métodos matemáticos, despertar su mente y enseñarles cierta cultura de estudios y aplicaciones de métodos matemáticos en su trabajo profesional.




ÁREA: MATEMÁTICAS
ISBN 968-18-6366-6




e-mail: limusa@noriega.com.mx
www.noriega.com.mx